TRANSACTIONS
OF THE
AMERICAN INSTITUTE
OF
ELECTRICAL ENGINEERS

VOLUME 73
1954

PART II. APPLICATIONS AND INDUSTRY

PUBLISHED BY THE
AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
33 WEST 39TH STREET
NEW YORK 18, N. Y., U. S. A.

Copyright 1954 by the American Institute of Electrical Engineers
Printed in U. S. A. by The Mack Printing Company, Easton, Pa.

Preface

The AIEE *Transactions* for 1954 (volume 73) is published in three parts: Part I. Communication and Electronics; Part II. Applications and Industry; and Part III. Power Apparatus and Systems. (Part III is divided into two parts: III-A, consisting of pages 1–832, and III-B, consisting of pages 833–1780. The Index for both parts appears at the end of III-B.) The papers in each of the three parts are classified according to subject matter as follows:

Part I. Communication and Electronics

Communication Switching Systems
Radio Communications Systems
Special Communications Applications
Telegraph Systems
Television and Aural Broadcasting Systems
Wire Communications Systems
Basic Sciences
Computing Devices
Electrical Techniques in Medicine and Biology
Electronic Power Converters
Electronics
Instruments and Measurements
Magnetic Amplifiers
Metallic Rectifiers
Nucleonics

Part II. Applications and Industry

Air Transportation
Domestic and Commercial Applications
Land Transportation
Marine Transportation
Production and Application of Light
Chemical, Electrochemical and Electrothermal Applications
Electric Heating
Electric Welding
Feedback Control Systems
General Industry Applications
Industrial Control
Industrial Power Systems
Mining and Metal Industry

Part III. Power Apparatus and Systems

Carrier Current
Insulated Conductors
Power Generation
Protective Devices
Relays
Rotating Machinery
Substations
Switchgear
System Engineering
Transformers
Transmission and Distribution

Education
Safety
Research

Each part has been indexed separately in the back of that particular part. The three parts are not cross-referenced; hence the user should determine first whether the subject matter of the paper desired is in the field of communication and electronics, applications and industry, or power apparatus and systems. Papers are listed in the subject index under several key words in the titles. The original numbers assigned to the papers are given in the author index. Volume 73 contains the technical papers and related discussions presented at these meetings:

1. Winter General Meeting, New York, N. Y., January 18–22, 1954.
2. North Eastern District Meeting, Schenectady, N. Y., May 5–7, 1954.
3. Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954.
4. Middle Eastern District Meeting, Reading, Pa., October 5–7, 1954.
5. Fall General Meeting, Chicago, Ill., October 11–15, 1954.

Statements and opinions given in the papers and discussions published in *Transactions* are the expressions of the contributors for which the American Institute of Electrical Engineers assumes no responsibility.

Applications and Industry Index for 1954

1. Technical Subject Index

A-C Generators, Characteristics of Aircraft. Stratton, Matsch....165–9
A-C Generators, Transient Characteristics of Aircraft. Holloway....187–90
A-C Motors and Controls in Refineries and Pipe Lines, Application of Large. Oscarson....296–301
A-C Power System at the Fairless Works. Watkins, Derr, Fountain, Squires....343–52
A-C Systems, A Method of Calculating Current-Limiter and Fuse-Clearing Times in. Caldwell, Jensen....263–7; disc. 267
A-C Systems, Overvoltage Detection in Single and Multigenerator Aircraft. Tucker, Trbovich....420–4; disc. 424
Accurate Control of Relative Speed and Cut in a Continuous Process Line. Anger, Pettit....485–90; disc. 490
Acid Batteries, Some Discharge Characteristics of Lead. Hoxie....17–22
Acid Storage Batteries, Some Aspects of the Charge and Discharge Processes in Lead. Craig, Hamer....22–34
Additions to z-Transformation Theory for Sampled-Data Systems. Lago....403–08
Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled Aircraft Generators, Ten Years of Progress in Predicting the. Scott....455–60; disc. 460
Air-Borne Stabilized Camera Mount, Development of an. Miller, Alexander....184–6
Air Ionization as an Environment Factor. Beckett....161–5
(Air Transportation) A High-Power Servo Analyzer. Gorrill, Walley....338–43
(Air Transportation) A Method of Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems. Caldwell, Jensen....263–7; disc. 267
(Air Transportation) A Self-Excited Induction Generator with Regulated Voltage. McConnell....288–94; disc. 294
(Air Transportation) Calculations on Voltage Unbalance for 3-Phase Synchronous Systems. Wilson, Gardner....426–37; disc. 437
(Air Transportation) Carbon-Pile Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation. Richards....357–63
(Air Transportation) Correlation of Rating Data for Rotating Electric Machinery. Naumann....443–51; disc. 451
(Air Transportation) Development of a Large Ratio-Release Circuit-Breaker Mechanism. Scarcelli, Steiner....512–17; disc. 517
(Air Transportation) Environmental Testing of Small Electronic Components at High Sound Intensities. Mintz, Levine....503–09; disc. 509
(Air Transportation) Evaluation of Designs for Intermittently Heated Surfaces. Dahm, Holloway....198–205
(Air Transportation) Performance of a Constant-Speed Drive. Giloy....179–84
(Air Transportation) The Nature of Voltage Ripple on D-C Generators. Collamore....390–3
Aircraft A-C Generators, Characteristics of. Stratton, Matsch....165–9
Aircraft A-C Generators, Considerations Applicable to Automatic Paralleling of. Powell, Giloy....258–63; disc. 263
Aircraft A-C Generators, Transient Characteristics of. Holloway....187–90
Aircraft A-C Systems, Overvoltage Detection in Single and Multigenerator. Tucker, Trbovich....420–4; disc. 424
Aircraft Electric Power Systems, Economic Factors for. Bergslien, Stratton, Finison....270–9; disc. 278
Aircraft Electric System, Study of Transformerless Rectified Higher Voltage D-C. Dallas, Reising....253–7; disc. 257
Aircraft Electric Systems, A Weight Analysis of Modern. Garbarino, Hawkes, Granath....463–8; disc. 468
Aircraft Electric Systems, Sensitivity Requirements of Reactive Load Division Circuits in. Sherrard....279–86; disc. 286
Aircraft Equipment, Reliability Improvement—A Plan to Achieve and Measure Reliability of. Luckman....378–83
Aircraft Generators, Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled. Scott....455–60; disc. 460
Aircraft Generators, Thermal Developments in. Martin....493–502; disc. 502
Aircraft Inverters, High Altitude. Franklin....363–72
Aircraft Power Distribution Circuits, Experimental Determination of 400-Cycle Impedance of Wire in. Andrew....469–77; disc. 477
Aircraft Switch Testing, Comments on. Stuelpnagel, Dallas....301–05
Aircraft Tachometer Indicator—An Analysis of Design Factors Affecting Starting Performance. Akeley....398–403
Aircraft Windshields Heated by Means of Transparent Conductive Films. Ward....408–16
Altitude Aircraft Inverters, High. Franklin....363–72
Analysis of Automatic Control Systems with Nonlinear Gain Elements, Phase-Plane. Kalman....383–90
Analysis of Design Factors Affecting Starting Performance, Aircraft Tachometer Indicator—An. Akeley....398–403
Analysis of Modern Aircraft Electric Systems, A Weight. Garbarino, Hawkes, Granath....463–8; disc. 468
Analysis of Sampled-Data Control Systems, Transient. Johnson, Lindorff....147–53
Analyzer, A High-Power Servo. Gorrill, Walley....338–43
Application of Large A-C Motors and Controls in Refineries and Pipe Lines. Oscarson....296–301
Application of Short-Time Memory Devices to Compensator Design, The. Ford, Calvert....88–93; disc. 93
Automatic Control Systems with Nonlinear Gain Elements, Phase-Plane Analysis of. Kalman....383–90
Automatic Paralleling of Aircraft A-C Generators, Considerations Applicable to. Powell, Giloy....258–63; disc. 263
"Automation"—Electrolytic Tinning, Trends in. Gravenstreter, Layton....97–100; disc. 100

B

Batteries, Some Aspects of the Charge and Discharge Processes in Lead-Acid Storage. Craig, Hamer....22–34
Batteries, Some Discharge Characteristics of Lead Acid. Hoxie....17–22
Blast-Cooled Aircraft Generators, Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of. Scott....455–60; disc. 460
Blooming Mill, Fairless Works—Electric Equipment for Slabbing Mill and. Wright, Kincaid....141–3; disc. 143
Bus Bars Under Unbalanced Load Conditions, Current Distribution in Paired-Phase. Cataldo, Shackman....438–42; disc. 442

C

Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems, A Method of. Caldwell, Jensen....263–7; disc. 267
Calculations on Voltage Unbalance for 3-Phase Synchronous Systems. Wilson, Gardner....426–37; disc. 437
Camera Mount, Development of an Airborne Stabilized. Miller, Alexander....184–6
Carbon-Pile Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation. Richards....357–63
Cell Lines, Magnetic Amplifiers in Metering Direct Current on Electrolytic. Downing....93–6; disc. 96
Characteristics of Aircraft A-C Generators. Stratton, Matsch....165–9
Characteristics of Aircraft A-C Generators, Transient. Holloway....187–90
Characteristics of Lead Acid Batteries, Some Discharge. Hoxie....17–22
Charge and Discharge Processes in Lead-Acid Storage Batteries, Some Aspects of the. Craig, Hamer....22–34
(Chemical) Some Discharge Characteristics of Lead Acid Batteries. Hoxie....17–22
Circuit-Breaker Mechanism, Development of a Large Ratio-Release. Scarcelli, Steiner....512–17; disc. 517
Circuit Design, Some Fundamentals of Equipment-Grounding. Kaufmann....227–31; disc. 231
Circuit Systems, Iron Conduit Impedance Effects in Ground. Bisson, Rochau....104–06; disc. 107
Circuits in Aircraft Electric Systems, Sensitivity Requirements of Reactive Load Division. Sherrard....279–86; disc. 286
Closed-Loop Frequency Response of Feedback Control Systems, An Extension of the Root Locus Method to Obtain. Jackson....176–9
Closed-Loop Frequency Response of Nonlinear Systems, A Generalized Method for Determining the. Prince....217–24; disc. 224
Comments on Aircraft Switch Testing. Stuelpnagel, Dallas....301–05
Commercial Buildings, Progress in Power System Engineering for. Kurt, Beeman....313–20; disc. 335
Commercial Buildings, The Relative Feasibility of 460-Volt or 208-Volt Service in. Barnett, Zimmerman, Lokay....306–13; disc. 337
(Commercial) Service Voltage Spread and Its Effect on Utilisation Equipment. Barnett, Lawrence....320–5; disc. 336
Compensation of Saturating Servomechanisms, Linear. Burnett, Kendall....6–10
Compensator Design, The Application of Short-Time Memory Devices to. Ford, Calvert....88–93; disc. 93
Components at High Sound Intensities, Environmental Testing of Small Electronic. Mintz, Levine....503–09; disc. 509
Conductive Films, Aircraft Windshields Heated by Means of Transparent. Ward....408–16
Conduit Impedance Effects in Ground Circuit Systems, Iron. Bisson, Rochau....104–06; disc. 107
Considerations Applicable to Automatic Paralleling of Aircraft A-C Generators. Powell, Giloy....258–63; disc. 263
Considerations in Applying Rectifiers as a Power Supply for Hot Strip Mills. Zins, Cham....65–9; disc. 69
Constant-Speed Drive, Performance of a. Giloy....179–84
Contactor Servomechanisms, An Investigation of the Switching Criteria for Higher Order. Bogner, Kazda....118–26; disc. 126
Contactor Servomechanisms Employing Sampled Data. Chow....51–62; disc. 62
Control for Diesel-Electric Locomotives, Static Excitation. McElhenny, Smith....158–60
Control of Radio-Frequency Generators, Magnetic Amplifier. Mohr, Lee....374–8
Control of Relative Speed and Cut in a Continuous Process Line, Accurate. Anger, Pettit....485–90; disc. 490
Control Systems, Transient Analysis of Sampled-Data. Johnson, Lindorff....147–53
Control Systems with Nonlinear Gain Elements, Phase-Plane Analysis of Automatic. Kalman....383–90
Co-ordinated Fuse Protection for Low-Voltage Distribution Systems in Industrial Plants. Lebens....77–80; disc. 80
Co-ordination Considerations, Rectifier Motive Power—Inductive. King, Gordon, Hibbard....107–14; disc. 114
Correlation of Rating Data for Rotating Electric Machinery. Naumann....443–51; disc. 451
Criteria for Higher Order Contactor Servomechanisms, An Investigation of the Switching. Bogner, Kazda....118–26; disc. 126
Criteria for Servomechanism Performance, The Influence of Time Scale and Gain on. Graham, Lathrop....153–7; disc. 157
Current Distribution in Paired-Phase Bus Bars Under Unbalanced Load Conditions. Cataldo, Shackman....438–42; disc. 442
Current-Limiter and Fuse-Clearing Times in A-C Systems, A Method of Calculating. Caldwell, Jensen....263–7; disc. 267
Cut in a Continuous Process Line, Accurate Control of Relative Speed and. Anger, Pettit....485–90; disc. 490

D

D-C Aircraft Electric System, Study of Transformerless Rectified Higher Voltage. Dallas, Reising..........253–7; disc. 257
D-C Diesel-Electric Ship-Propulsion Systems, Series-Versus Parallel-Connected Generators for Multiple-Engine. Wasmund..........135–8; disc. 138
D-C Generators, The Nature of Voltage Ripple on. Collamore..........390–3
D-C Positional Servomechanisms, A Magnetic Tape Memory for. DeBarber..........372–4
Data Control Systems, Transient Analysis of Sampled-. Johnson, Lindorff..........147–53
Data Feedback Systems, The Design of Sampled-. Lago, Truxall..........247–53; disc. 253
Data for Rotating Electric Machinery, Correlation of Rating. Naumann..........443–51; disc. 451
Data Processing Techniques for Feedback Control Systems, Sampled- Bergen, Ragazzini..........236–44; disc. 244
Data Systems, Additions to z-Transformation Theory for Sampled-. Lago..........403–08
Derivative and Integral Control, The Transient Performance of Servomechanisms with. Lathrop, Graham..........10–17
Design of Sampled-Data Feedback Systems, The. Lago, Truxall..........247–53; disc. 253
Designs for Intermittently Heated Surfaces, Evaluation of. Dahm, Holloway..........198–205
Development of a Large Ratio-Release Circuit-Breaker Mechanism. Scarcelli, Steiner...512–17; disc. 517
Development of an Air-Borne Stabilized Camera Mount. Miller, Alexander..........184–6
Developments in Aircraft Generators, Thermal. Martin..........493–502; disc. 502
Developments in the Theory and Design of Electric Spark Machine Tools, Recent. Williams, Woodford, Smith..........83–6; disc. 87
Devices to Compensator Design, The Application of Short-Time Memory. Ford, Calvert..........88–93; disc. 93
Diesel-Electric Locomotive?, A Reappraisal of the Economics of Railway Electrification: How, When, and Where Can It Compete with the. Brown, Kimball..........35–45; disc. 45
Diesel-Electric Locomotives, A World-Wide Standard Traction Motor for. Baldwin..........144–7
Diesel-Electric Locomotive Wheel Slipping—Causes, Effects, and Methods of Control. Ford..........478–80; disc. 480
Diesel-Electric Locomotives, Gearing for. Bevan..........510–12; disc. 512
Diesel-Electric Locomotives, Maintenance Testing of Insulation Resistance on. Kelley..........452–4; disc. 454
Diesel-Electric Locomotives Results in Design Changes, Operating Experience in. Legg..461–3; disc. 463
Diesel-Electric Locomotives, Static Excitation Control for. McElhenny, Smith..........158–60
Diesel-Electric Motors and Generators, Fundamentals of Flashing of. Atwell..........70–6; disc. 76
Diesel-Electric Ship-Propulsion Systems, Series-Versus Parallel-Connected Generators for Multiple-Engine D-C. Wasmund..........135–8; disc. 138
Direct Current on Electrolytic Cell Lines, Magnetic Amplifiers in Metering. Downing..93–6; disc. 96
Discharge Characteristics of Lead Acid Batteries, Some. Hoxie..........17–22
Discharge Processes in Lead-Acid Storage Batteries, Some Aspects of the Charge and. Craig, Hamer..........22–34
Distribution Circuits, Experimental Determination of 400-Cycle Impedance of Wire in Aircraft Power. Andrew..........469–77; disc. 477
Distribution in Paired-Phase Bus Bars Under Unbalanced Load Conditions, Current. Cataldo, Shackman..........438–42; disc. 442
Distribution Systems in Industrial Plants, Co-ordinated Fuse Protection for Low-Voltage. Lebens..........77–80; disc. 80
Distribution Voltage for Metropolitan Areas, Symposium on Higher..........306–35; disc. 335
Division Circuits in Aircraft Electric Systems, Sensitivity Requirements of Reactive Load. Sherrard..........279–86; disc. 286
Drive, Performance of a Constant-Speed. Giloy..........179–84

E

Economic Factors for Aircraft Electric Power Systems. Bergslien, Stratton, Finison..........270–8; disc. 278
Economics of Railway Electrification: How, When, and Where Can It Compete with the Diesel-Electric Locomotive, A Reappraisal of the. Brown, Kimball..........35–45; disc. 45
Electric Equipment for Slabbing Mill and Blooming Mill, Fairless Works—. Wright, Kincaid..........141–3; disc. 143
(Electric Heating) Induction Heating Steel with 60 Cycles. Kramer..........353–5; disc. 356
(Electric Heating) Magnetic Amplifier Control of Radio-Frequency Generators. Mohr, Lee..........374–8
Electric Machinery, Correlation of Rating Data for Rotating. Naumann..........443–51; disc. 451
Electric Power Systems, Economic Factors for Aircraft. Bergslien, Stratton, Finison..........270–8; disc. 278
Electric Spark Machine Tools, Recent Developments in the Theory and Design of. Williams, Woodford, Smith..........83–6; disc. 87
Electric System for Monorail Rapid Transit, An. Anson, Kimball..........225–6; disc. 226
Electrification: How, When, and Where Can It Compete with the Diesel-Electric Locomotive?, A Reappraisal of the Economics of Railway. Brown, Kimball..........35–45; disc. 45
Electrolytic and Electrothermal Applications, A New High-Current Switch for. Graybill..........1–6
Electrolytic Cell Lines, Magnetic Amplifiers in Metering Direct Current on. Downing..........93–6; disc. 96
Electrolytic Tinning, Trends in "Automation"—. Gravenstreter, Layton..........97–100; disc. 100
Electronic Components at High Sound Intensities, Environmental Testing of Small. Mintz, Levine..........503–09; disc. 509
Electrothermal Applications, A New High-Current Switch for Electrolytic and. Graybill..........1–6
Engineering for Commercial Buildings, Progress in Power System. Kurt, Beeman..313–20; disc. 335
Environment Factor, Air Ionization as an. Beckett..........161–5
Environmental Testing of Small Electronic Components at High Sound Intensities. Mintz, Levine..........503–09; disc. 509
Equipment, High-Speed Rapid Transit. Vouch..........267–70; disc. 270
Equipment for Slabbing Mill and Blooming Mill, Fairless Works—Electric. Wright, Kincaid..........141–3; disc. 143
Equipment-Grounding Circuit Design, Some Fundamentals of. Kaufmann..........227–31; disc. 231
Equipment, Reliability Improvement—A Plan to Achieve and Measure Reliability of Aircraft. Luckman..........378–83
Evaluating the Effect of Nonlinearity in a 2-Phase Servomotor. Stein, Thaler..........518–21; disc. 521
Evaluation of Designs for Intermittently Heated Surfaces. Dahm, Holloway..........198–205
Excitation Control for Diesel-Electric Locomotives Static. McElhenny, Smith..........158–60
Experimental Determination of 400-Cycle Impedance of Wire in Aircraft Power Distribution Circuits. Andrew..........469–77; disc. 477
Extension of the Root Locus Method to Obtain Closed-Loop Frequency Response of Feedback Control Systems, An. Jackson..........176–9

F

Fairless Works, A-C Power System at the. Watkins, Derr, Fountain, Squires..........343–52
Fairless Works—Electric Equipment for Slabbing Mill and Blooming Mill. Wright, Kincaid..........141–3; disc. 143
(Feedback Control) A Generalized Method for Determining the Closed-Loop Frequency Response of Nonlinear Systems. Prince.....217–24; disc. 224
(Feedback Control) A Magnetic Tape Memory for D-C Positional Servomechanisms. DeBarber....372–4
(Feedback Control) Additions to z-Transformation Theory for Sampled-Data Systems. Lago..........403–08
(Feedback Control) Evaluating the Effect of Nonlinearity in a 2-Phase Servomotor. Stein, Thaler..........518–21; disc. 521
(Feedback Control) Jitter in Instrument Servos. Hovious..........393–8
(Feedback Control) Measurement of Some Nonlinearities in Servomechanisms. Gehmlich, VanValkenburg..........232–5; disc. 235
(Feedback Control) Phase-Plane Analysis of Automatic Control Systems with Nonlinear Gain Elements. Kalman..........383–90
(Feedback Control) Self-Oscillation Method for Measuring Transfer Functions. Clegg, Harris.....169–70
Feedback Control Systems, An Extension of the Root Locus Method to Obtain Closed-Loop Frequency Response of. Jackson..........176–9
Feedback Control Systems, Sampled-Data Processing Techniques for. Bergen, Ragazzini..........236–44; disc. 244
(Feedback Control) The Application of Short-Time Memory Devices to Compensator Design. Ford, Calvert..........88–93; disc. 93
(Feedback Control) The Maximum Response Ratio of Linear Systems. Pfeiffer..........480–3; disc. 483
Feedback Systems, The Design of Sampled-Data. Lago, Truxall..........247–53; disc. 253
Films, Aircraft Windshields Heated by Means of Transparent Conductive. Ward..........408–16
Flashing of Diesel-Electric Motors and Generators, Fundamentals of. Atwell..........70–6; disc. 76
400-Cycle Impedance of Wire in Aircraft Power Distribution Circuits, Experimental Determination of. Andrew..........469–77; disc. 477
480 Wye/277-Volt Power System in Telephone Building at Menands, N.Y. Brereton, Donnelly..........325–35; disc. 338
460-Volt or 208-Volt Service in Commercial Buildings, The Relative Feasibility of. Barnett, Zimmerman, Lokay..........306–13; disc. 337
Frequency Response of Feedback Control Systems, An Extension of the Root Locus Method to Obtain Closed-Loop. Jackson..........176–9
Frequency Response of Nonlinear Systems, A Generalized Method for Determining the Closed-Loop. Prince..........217–24; disc. 224
Functions, Self-Oscillation Method for Measuring Transfer. Clegg, Harris..........169–70
Fundamentals of Equipment-Grounding Circuit Design, Some. Kaufmann..........227–31; disc. 231
Fundamentals of Flashing of Diesel-Electric Motors and Generators. Atwell..........70–6; disc. 76
Fuse-Clearing Times in A-C Systems, A Method of Calculating Current-Limiter and. Caldwell, Jensen..........263–7; disc. 267
Fuse Protection for Low-Voltage Distribution Systems in Industrial Plants, Co-ordinated. Lebens..........77–80; disc. 80

G

Gain Elements, Phase-Plane Analysis of Automatic Control Systems with Nonlinear. Kalman.....383–90
Gain of a Servomechanism for a Specified Maximum Modulus Less Than Unity, A Graphical Procedure for Determining the. Higgins..........101–03; disc. 103
Gain on Criteria for Servomechanism Performance, The Influence of Time Scale and. Graham, Lathrop..........153–7; disc. 157
Gearing for Diesel-Electric Locomotives. Bevan..........510–12; disc. 512
Generalized Method for Determining the Closed-Loop Frequency Response of Nonlinear Systems, A. Prince..........217–24; disc. 224
Generator with Regulated Voltage. A Self-Excited Induction. McConnell..........288–94; disc. 294
Generators, Characteristics of Aircraft A-C. Stratton, Matsch..........165–9
Generators, Considerations Applicable to Automatic Paralleling of Aircraft A-C. Powell, Giloy..........258–63; disc. 263
Generators for Multiple-Engine D-C Diesel-Electric Ship-Propulsion Systems, Series- Versus Parallel-Connected. Wasmund..........135–8; disc. 138
Generators, Fundamentals of Flashing of Diesel-Electric Motors and. Atwell..........70–6; disc. 76
Generators, Magnetic Amplifier Control of Radio-Frequency. Mohr, Lee..........374–8
Generators, Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled Aircraft. Scott..........455–60; disc. 460
Generators, The Nature of Voltage Ripple on D-C. Collamore..........390–3
Generators, Thermal Developments in Aircraft. Martin..........493–502; disc. 502
Generators, Transient Characteristics of Aircraft A-C. Holloway..........187–90
Graphical Procedure for Determining the Gain of a Servomechanism for a Specified Maximum Modulus Less Than Unity, A. Higgins..........101–03; disc. 103
Ground Circuit Systems, Iron Conduit Impedance Effects in. Bisson, Rochau......104–06; disc. 107
Grounding Circuit Design, Some Fundamentals of Equipment-. Kaufmann..........227–31; disc. 231

H

Heated Surfaces, Evaluation of Designs for Intermittently. Dahm, Holloway..........198–205
Heating Steel with 60 Cycles, Induction. Kramer..........353–5; disc. 356
High Altitude Aircraft Inverters. Franklin.....363–72
High-Current Switch for Electrolytic and Electrothermal Applications, A New. Graybill..........1–6
High-Power Servo Analyzer, A. Gorrill, Walley.338–43
High-Speed Rapid Transit Equipment. Vouch..........267–70; disc. 270
Hot Strip Mills, Considerations in Applying Rectifiers as a Power Supply for. Zins, Cham..........65–9; disc. 69

I

Ignitron Rectifier Locomotives on the Pennsylvania Railroad, Some Application Phases of the. Brown..........128–33; disc. 133
Impedance Effects in Ground Circuit Systems, Iron Conduit. Bimon, Rochau..........104–06; disc. 107
Impedance of Wire in Aircraft Power Distribution Circuits, Experimental Determination of 400-Cycle. Andrew..........469–77; disc. 477
Improvement—A Plan to Achieve and Measure Reliability of Aircraft Equipment, Reliability. Luckman..........378–83
Indicator—An Analysis of Design Factors Affecting Starting Performance, Aircraft Tachometer. Akeley..........398–403
Induction Generator with Regulated Voltage, A Self-Excited. McConnell..........288–94; disc. 294
Induction Heating Steel with 60 Cycles. Kramer..........353–5; disc. 356
Inductive Co-ordination Considerations, Rectifier Motive Power—. King, Gordon, Hibbard..........107–14; disc. 114
(Industrial Control) Accurate Control of Relative Speed and Cut in a Continuous Process Line. Anger, Pettit..........485–90; disc. 490
Industrial Plants, Co-ordinated Fuse Protection for Low-Voltage Distribution Systems in. Lebens..........77–80; disc. 80
(Industrial Power Systems) A-C Power System at the Fairless Works. Watkins, Derr, Fountain, Squires..........343–52
(Industrial Power Systems) Current Distribution in Paired-Phase Bus Bars Under Unbalanced Load Conditions. Cataldo, Shackman..........438–42; disc. 442
(Industrial Power Systems) 480-Wye/277-Volt Power System in Telephone Building at Menands, N.Y. Brereton, Donnelly..........325–35; disc. 338
(Industrial Power Systems) Progress in Power System Engineering for Commercial Buildings. Kurt, Beeman..........313–20; disc. 335
(Industrial Power Systems) Service Voltage Spread and Its Effect on Utilization Equipment. Barnett, Lawrence..........320–5; disc. 336
(Industrial Power Systems) The Relative Feasibility of 460-Volt or 208-Volt Service in Commercial Buildings. Barnett, Zimmerman, Lokay..........306–13; disc. 337
Influence of Time Scale and Gain on Criteria for Servomechanism Performance, The. Graham, Lathrop..........153–7; disc. 157
Instrument Servos, Jitter in. Hovious..........393–8
Insultation Resistance on Diesel-Electric Locomotives, Maintenance Testing of. Kelley..452–4; disc. 454
Integral Control, The Transient Performance of Servomechanisms with Derivative and. Lathrop, Graham..........10–17
Intermittently Heated Surfaces, Evaluation of Designs for. Dahm, Holloway..........198–205
Inverters, High Altitude Aircraft. Franklin....363–72
Investigation of the Switching Criteria for Higher Order Contactor Servomechanisms, An. Bogner, Kazda..........118–26; disc. 126
Ionization as an Environment Factor, Air. Beckett..........161–5
Iron Conduit Impedance Effects in Ground Circuit Systems. Bimon, Rochau..........104–06; disc. 107

J

Jitter in Instrument Servos. Hovious..........393–8

L

(Land Transportation) A Reappraisal of the Economics of Railway Electrification: How, When, and Where Can It Compete with the Diesel-Electric Locomotive? Brown, Kimball..........35–45; disc. 45
(Land Transportation) A World-Wide Standard Traction Motor for Diesel-Electric Locomotives. Baldwin..........144–7
(Land Transportation) An Electric System for Monorail Rapid Transit. Anson, Kimball..225–6; disc. 226
(Land Transportation) Diesel-Electric Locomotive Wheel Slipping—Causes, Effects, and Methods of Control. Fort..........478–80; disc. 480
(Land Transportation) Fundamentals of Flashing of Diesel-Electric Motors and Generators. Atwell..........70–6; disc. 76
(Land Transportation) Gearing for Diesel-Electric Locomotives. Bevan..........510–12; disc. 512
(Land Transportation) High-Speed Rapid Transit Equipment. Vouch..........267–70; disc. 270
(Land Transportation) Maintenance Testing of Insulation Resistance on Diesel-Electric Locomotives Kelley..........452–4; disc. 454
(Land Transportation) Operating Experience in Diesel-Electric Locomotives Results in Design Changes. Legg..........461–3; disc. 463
(Land Transportation) Rebuilding San Francisco's Transit System. ReQua..........171–6
(Land Transportation) Rectifier Motive Power—Inductive Co-ordination Considerations. King, Gordon, Hibbard..........107–14; disc. 114
(Land Transportation) Some Application Phases of the Ignitron Rectifier Locomotives on the Pennsylvania Railroad. Brown..........128–33; disc. 133
(Land Transportation) Static Excitation Control for Diesel-Electric Locomotives. McElhenny, Smith..........158–60
(Land Transportation) The Toronto Subway. Inglis..........416–20; disc. 420
Lead Acid Batteries, Some Discharge Characteristics of. Hoxie..........17–22
Lead-Acid Storage Batteries, Some Aspects of the Charge and Discharge Processes in. Craig, Hamer..........22–34
Linear Compensation of Saturating Servomechanisms. Burnett, Kendall..........6–10
Linear Systems, The Maximum Response Ratio of. Pfeiffer..........480–3; disc. 483
Load Conditions, Current Distribution in Paired-Phase Bus Bars Under Unbalanced. Cataldo, Shackman..........438–42; disc. 442
Load Division Circuits in Aircraft Electric Systems, Sensitivity Requirements of Reactive. Sherrard..........279–86; disc. 286
Locomotive?, A Reappraisal of the Economics of Railway Electrification: How, When, and Where Can It Compete with the Diesel-Electric. Brown, Kimball..........35–45; disc. 45
Loromotive Wheel Slipping—Causes, Effects, and Methods of Control, Diesel-Electric. Fort..........478–80; disc. 480
Locomotives, A World-Wide Standard Traction Motor for Diesel-Electric. Baldwin..........144–7
Locomotives, Gearing for Diesel-Electric. Bevan..........510–12; disc. 512
Locomotives, Maintenance Testing of Insulation Resistance on Diesel-Electric. Kelley..452–4; disc. 454
Locomotives on the Pennsylvania Railroad. Some Application Phases of the Ignitron Rectifier. Brown..........128–33; disc. 133
Locomotives Results in Design Changes, Operating Experience in Diesel-Electric. Legg..........461–3; disc. 463
Locomotives, Static Excitation Control for Diesel-Electric. McElhenny, Smith..........158–60
Locus Method to Obtain Closed-Loop Frequency Response of Feedback Control Systems, An Extension of the Root. Jackson..........176–9
Low-Voltage Distribution Systems in Industrial Plants, Co-ordinated Fuse Protection for. Lebens..........77–80; disc. 80

M

Machine Tools, Recent Developments in the Theory and Design of Electric Spark. Williams, Woodford, Smith..........83–6; disc. 87
Magnetic Amplifier Control of Radio-Frequency Generators. Mohr, Lee..........374–8
Magnetic Amplifiers in Metering Direct Current on Electrolytic Cell Lines. Downing..93–6; disc. 96
Magnetic Tape Memory for D-C Positional Servomechanisms, A. DeBarber..........372–4
Maintenance Testing of Insulation Resistance on Diesel-Electric Locomotives. Kelley....452–4; disc. 454
(Marine Transportation) Series Versus Parallel-Connected Generators for Multiple-Engine D-C Diesel-Electric Ship-Propulsion Systems. Wasmund..........135–8; disc. 138
Maximum Modulus Less Than Unity, A Graphical Procedure for Determining the Gain of a Servomechanism for a Specified. Higgins..101–03; disc. 103
Maximum Response Ratio of Linear Systems, The. Pfeiffer..........480–3; disc. 483
Measurement of Some Nonlinearities in Servomechanisms. Gehmlich, VanValkenburg..........232–5; disc. 235
Measuring Transfer Functions, Self-Oscillation Method for. Clegg, Harris..........169–70
Mechanism, Development of a Large Ratio-Release Circuit-Breaker. Scarcelli, Steiner..512–17; disc. 517
Memory Devices to Compensator Design, The Application of Short-Time. Ford, Calvert..........88–93; disc. 93
Memory for D-C Positional Servomechanisms, A Magnetic Tape. DeBarber..........372–4
Menands, N. Y. 480-Wye/277-Volt Power System in Telephone Building at. Brereton, Donnelly..........325–35; disc. 338
(Metal) Fairless Works—Electric Equipment for Slabbing Mill and Blooming Mill. Wright, Kincaid..........141–3; disc. 143
Metering Direct Current on Electrolytic Cell Lines, Magnetic Amplifiers in. Downing......93–6; disc. 96
Method for Determining the Closed-Loop Frequency Response of Nonlinear Systems, A Generalized. Prince..........217–24; disc. 224
Method for Measuring Transfer Functions, Self-Oscillation. Clegg, Harris..........169–70
Method of Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems, A. Caldwell, Jensen..........263–7; disc. 267
Method to Obtain Closed-Loop Frequency Response of Feedback Control Systems, An Extension of the Root Locus. Jackson..........176–9
Metropolitan Areas, Symposium on Higher Distribution Voltage for..........306–35; disc. 335
Mill and Blooming Mill, Fairless Works—Electric Equipment for Slabbing. Wright, Kincaid..........141–3; disc. 143
Mills, Considerations in Applying Rectifiers as a Power Supply for Hot Strip. Zins, Cham..........65–9; disc. 69
Mills with Large Utility Power Systems, Problems Relating to Interconnections of Large Pulp and Paper. Rose, Springer..........205–14; disc. 214
(Mining) Trends in "Automation"—Electrolytic Tinning. Gravenstreter, Layton...97–100; disc. 100
Modulus Less Than Unity, A Graphical Procedure for Determining the Gain of a Servomechanism for a Specified Maximum. Higgins...101–03; disc. 103
Monorail Rapid Transit, An Electric System for. Anson, Kimball..........225–6; disc. 226
Motive Power—Inductive Co-ordination Considerations, Rectifier. King, Gordon, Hibbard..........107–14; disc. 114
Motor for Diesel-Electric Locomotives, A World-Wide Standard Traction. Baldwin..........144–7
Motors and Controls in Refineries and Pipe Lines, Application of Large A-C. Oscarson..........296–301
Motors and Generators, Fundamentals of Flashing of Diesel-Electric. Atwell..........70–6; disc. 76
Mount, Development of an Air-Borne Stabilized Camera. Miller, Alexander..........184–6
Multigenerator Aircraft A-C Systems, Overvoltage Detection in Single and. Tucker, Trbovich..........420–4; disc. 424
Multiple-Engine D-C Diesel-Electric Ship-Propulsion Systems, Series- Versus Parallel-Connected Generators for. Wasmund..........135–8; disc. 138

N

Nature of Voltage Ripple on D-C Generators, The. Collamore..........390–3
Nonlinear Gain Elements, Phase-Plane Analysis of Automatic Control Systems with. Kalman.....383–90
Nonlinear Systems, A Generalized Method for Determining the Closed-Loop Frequency Response of. Prince..........217–24; disc. 224
Nonlinearities in Servomechanisms, Measurement of Some. Gehmlich, VanValkenburg..........232–5; disc. 235
Nonlinearity in a 2-Phase Servomotor, Evaluating the Effect of. Stein, Thaler..........518–21; disc. 521

O

Operating Experience in Diesel-Electric Locomotives Results in Design Changes. Legg..........461–3; disc. 463
Operation of Pipe-Line Pumping Stations, Remote. Derr, Hyde..........190–8
Oscillation Method for Measuring Transfer Functions, Self-. Clegg, Harris..........169–70
Output Characteristics of Blast-Cooled Aircraft Generators, Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and. Scott..........455–60; disc. 460
Overvoltage Detection in Single and Multigenerator Aircraft A-C Systems. Tucker, Trbovich..........420–4; disc. 424

P

Paired-Phase Bus Bars Under Unbalanced Load Conditions, Current Distribution in. Cataldo, Shackman..........438–42; disc. 442
Paper Mills with Large Utility Power Systems, Problems Relating to Interconnections of Large Pulp and. Rose, Springer..........205–14; disc. 214
Parallel-Connected Generators for Multiple-Engine D-C Diesel-Electric Ship-Propulsion Systems, Series Versus. Wasmund..........135–8; disc. 138
Paralleling of Aircraft A-C Generators, Considerations Applicable to Automatic. Powell, Giloy..........258–63; disc. 263

Pennsylvania Railroad, Some Application Phases of the Ignitron Rectifier Locomotives on the. Brown.......128–33; disc. 133
Performance of a Constant-Speed Drive. Giloy..179–84
Performance of Servomechanisms with Derivative and Integral Control, The Transient. Lathrop, Graham.......10–17
(Petroleum Industry) Application of Large A-C Motors and Controls in Refineries and Pipe Lines. Oscarson.......296–301
Phase-Plane Analysis of Automatic Control Systems with Nonlinear Gain Elements. Kalman.......383–90
Pipe-Line Pumping Stations, Remote Operation of. Derr, Hyde.......190–8
Pipe Lines, Application of Large A-C Motors and Controls in Refineries and. Oscarson.......296–301
Positional Servomechanisms, A Magnetic Tape Memory for D-C. DeBarber.......372–4
Power Distribution Circuits, Experimental Determination of 400-Cycle Impedance of Wire in Aircraft. Andrew.......469–77; disc. 477
Power Supply for Hot Strip Mills, Considerations in Applying Rectifiers as a. Zins, Cham.......65–9; disc. 69
Power System at the Fairless Works, A-C. Watkins, Derr, Fountain, Squires.......343–52
Power System Engineering for Commercial Buildings, Progress in. Kurt, Beeman.......313–20; disc. 335
Power System in Telephone Building at Menands, N. Y., 480 Wye/277-Volt. Brereton, Donnelly.......325–35; disc. 338
Power Systems, Economic Factors for Aircraft Electric. Bergslien, Stratton, Finison.......270–8; disc. 278
Power Systems, Problems Relating to Interconnections of Large Pulp and Paper Mills with Large Utility. Rose, Springer.......205–14; disc. 214
Problems Relating to Interconnections of Large Pulp and Paper Mills with Large Utility Power Systems. Rose, Springer.......205–14; disc. 214
Process Line, Accurate Control of Relative Speed and Cut in a Continuous. Anger, Pettit.......485–90; disc. 490
Processes in Lead Acid Storage Batteries, Some Aspects of the Charge and Discharge. Craig, Hamer.22–34
Progress in Power System Engineering for Commercial Buildings. Kurt, Beeman.......313–20; disc. 335
Propulsion Systems, Series-Versus Parallel-Connected Generators for Multiple-Engine D-C Diesel-Electric Ship-. Wasmund.......135–8; disc. 138
Pulp and Paper Mills with Large Utility Power Systems, Problems Relating to Interconnections of Large. Rose, Springer.......205–14; disc. 214
Pumping Stations, Remote Operation of Pipe-Line. Derr, Hyde.......190–8

R

Radio Frequency Generators, Magnetic Amplifier Control of. Mohr, Lee.......374–8
Railroad, Some Application Phases of the Ignitron Rectifier Locomotives on the Pennsylvania. Brown.......128–33; disc. 133
Railway Electrification: How, When, and Where Can it Compete with the Diesel-Electric Locomotive?, A Reappraisal of the Economics of. Brown, Kimball.......35–45; disc. 45
Rapid Transit, An Electric System for Monorail. Anson, Kimball.......225–6; disc. 226
Rapid Transit Equipment, High-Speed. Vouch.......267–70; disc. 270
Rating Data for Rotating Electric Machinery, Correlation of. Naumann.......443–51; disc. 451
Ratio of Linear Systems, The Maximum Response. Pfeiffer.......480–3; disc. 483
Ratio-Release Circuit-Breaker Mechanism, Development of a Large. Scarcelli, Steiner.......512–17; disc. 517
Reactive Load Division Circuits in Aircraft Electric Systems, Sensitivity Requirements of. Sherrard.......279–86; disc. 286
Reappraisal of the Economics of Railway Electrification; How, When, and Where Can It Compete with the Diesel-Electric Locomotive, A. Brown, Kimball.......35–45; disc. 45
Rebuilding San Francisco's Transit System. ReQua.......171–6
Recent Developments in the Theory and Design of Electric Spark Machine Tools. Williams, Woodford, Smith.......83–6; disc. 87
Rectifier Locomotives on the Pennsylvania Railroad, Some Application Phases of the Ignitron. Brown.......128–33; disc. 133
Rectifier Motive-Power—Inductive Co-ordination Considerations. King, Gordon, Hibbard.......107–14; disc. 114
Rectifiers as a Power Supply for Hot Strip Mills, Considerations in Applying. Zins, Cham.......65–9; disc. 69
Refineries and Pipe Lines, Application of Large A-C Motors and Controls in. Oscarson.......296–301
Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation, Carbon-Pile. Richards.......357–63
Reliability Improvement—A Plan to Achieve and Measure Reliability of Aircraft Equipment. Luckman.......378–83
Remote Operation of Pipe-Line Pumping Stations. Derr, Hyde.......190–8
Response Ratio of Linear Systems, The Maximum. Pfeiffer.......480–3; disc. 483
Root Locus Method to Obtain Closed-Loop Frequency Response of Feedback Control Systems, An Extension of the. Jackson.......176–9
Rotating Electric Machinery, Correlation of Rating Data for. Naumann.......443–51; disc. 451

S

Sampled Data, Contactor Servomechanisms Employing. Chow.......51–62; disc. 62
Sampled-Data Control Systems, Transient Analysis of. Johnson, Lindorff.......147–53
Sampled-Data Feedback Systems, The Design of. Lago, Truxall.......247–53; disc. 253
Sampled-Data Processing Techniques for Feedback Control Systems. Bergen, Ragazzini.......236–44; disc. 244
Sampled-Data Systems, Additions to z-Transformation Theory for. Lago.......403–08
San Francisco's Transit System, Rebuilding. ReQua.......171–6
Saturating Servomechanisms, Linear Compensation of. Burnett, Kendall.......6–10
Scale and Gain on Criteria for Serfomechanism Performance, The Influence of Time. Graham, Lathrop.......153–7; disc. 157
Self-Excited Induction Generator with Regulated Voltage, A. McConnell.......288–94; disc. 294
Self-Oscillation Method for Measuring Transfer Functions. Clegg, Harris.......169–70
Sensitivity Requirements of Reactive Load Division Circuits in Aircraft Electric Systems. Sherrard.......279–86; disc. 286
Series-Versus Parallel-Connected Generators for Multiple-Engine D-C Diesel-Electric Ship-Propulsion Systems. Wasmund.......135–8; disc. 138
Service Voltage Spread and Its Effect on Utilization Equipment. Barnett, Lawrence..320–5; disc. 336
Servo Analyzer, A High-Power. Gorrill, Walley.......338–43
Servomechanism for a Specified Maximum Modulus Less Than Unity, A Graphical Procedure for Determining the Gain of a. Higgins..101–03; disc. 103
Servomechanism Performance, The Influence of Time Scale and Gain on Criteria for. Graham, Lathrop.......153–7; disc. 157
Servomechanisms, A Magnetic Tape Memory for D-C Positional. DeBarber.......372–4
Servomechanisms, An Investigation of the Switching Criteria for Higher Order Contactor. Bogner, Kazda.......118–26; disc. 126
Servomechanisms Employing Sampled Data, Contactor. Chow.......51–62; disc. 62
Servomechanisms, Linear Compensation of Saturating. Burnett, Kendall.......6–10
Servomechanisms, Measurement of Some Nonlinearities in. Gahmlich, VanValkenburg...232–5; disc. 235
Servomechanisms with Derivative and Integral Control, The Transient Performance of. Lathrop, Graham.......10–17
Servomotor, Evaluating the Effect of Nonlinearity in a 2-Phase. Stein, Thaler.......518–21; disc. 521
Servos, Jitter in Instrument. Hovious.......393–8
Ship-Propulsion Systems, Series-Versus Parallel-Connected Generators for Multiple-Engine D-C Diesel-Electric. Wasmund.......135–8; disc. 138
Single and Multigenerator Aircraft A-C Systems, Overvoltage Detection in. Tucker, Trbovich.......420–4; disc. 424
60 Cycles, Induction Heating Steel with. Kramer.......353–5; disc. 356
Slabbing Mill and Blooming Mill, Fairless Works—Electric Equipment for. Wright, Kincaid.......141–3; disc. 143
Some Application Phases of the Ignitron Rectifier Locomotives on the Pennsylvania Railroad. Brown.......128–33; disc. 133
Some Aspects of the Charge and Discharge Processes in Lead Acid Storage Batteries. Craig, Hamer.22–34
Some Discharge Characteristics of Lead Acid Batteries. Hoxie.......17–22
Sound Intensities, Environmental Testing of Small Electronic Components at High. Mintz, Levine.......503–09; disc. 509
Spark Machine Tools, Recent Developments in the Theory and Design of Electric. Williams, Woodford, Smith.......83–6; disc. 87
Speed and Cut in a Continuous Process Line, Accurate Control of Relative. Anger, Pettit.......485–90; disc. 490
Standard Traction Motor for Diesel-Electric Locomotives, A World-Wide. Baldwin.......144–7
Starting Performance, Aircraft Tachometer Indicator—An Analysis of Design Factors Affecting. Akeley.......398–403
Static Excitation Control for Diesel-Electric Locomotives McElhenny, Smith.......158–60
Steel with 60 Cycles, Induction Heating. Kramer.......353–5; disc. 356
Storage Batteries, Some Aspects of the Charge and Discharge Processes in Lead-Acid. Craig, Hamer.......22–34
Strip Mills, Considerations in Applying Rectifiers as a Power Supply for Hot. Zins, Cham..65–9; disc. 69
Study of Transformerless Rectified Higher Voltage D-C Aircraft Electric System. Dallas, Reising.......253–7; disc. 257
Subway, The Toronto. Inglis.......416–20; disc. 420
Surfaces, Evaluation of Designs for Intermittently Heated. Dahm, Holloway.......198–205
Switch for Electrolytic and Electrothermal Applications, A New High-Current. Graybill.......1–6
Switch Testing, Comments on Aircraft. Stuelpnagel, Dallas.......301–05
Switching Criteria for Higher Order Contactor Servomechanisms, An Investigation of the. Bogner, Kazda.......118–26; disc. 126
Symposium on Higher Distribution Voltage for Metropolitan Areas.......306–35; disc. 335
Synchronous Systems, Calculations on Voltage Unbalance for 3-Phase. Wilson, Gardner.......426–37; disc. 437

T

Tachometer Indicator—An Analysis of Design Factors Affecting Starting Performance, Aircraft. Akeley.......398–403
Tape Memory for D-C Positional Servomechanisms, A Magnetic. DeBarber.......372–4
Telephone Building at Menands, N. Y., 480-Wye/270-Volt Power System in. Brereton, Donnelly.......325–35; disc. 338
Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled Aircraft Generators. Scott.......455–60; disc. 460
Testing, Comments on Aircraft Switch. Stuelpnagel, Dallas.......301–05
Testing of Insulation Resistance on Diesel-Electric Locomotives, Maintenance. Kelley..452–4; disc. 454
Testing of Small Electronic Components at High Sound Intensities, Environmental. Mintz, Levine.......503–09; disc. 509
Theory and Design of Electric Spark Machine Tools, Recent Developments in the. Williams, Woodford, Smith.......83–6; disc. 87
Theory: Calibration, Adjustment, and Factors Affecting Its Operation, Carbon-Pile Regulator. Richards.......357–63
Theory for Sampled-Data Systems, Additions to z-Transformation. Lago.......403–08
Thermal Developments in Aircraft Generators. Martin.......493–502; disc. 502
Thermodynamic and Output Characteristics of Blast-Cooled Aircraft Generators, Ten Years of Progress in Predicting the Aerodynamic. Scott.......455–60; disc. 460
Three-Phase Synchronous Systems, Calculations on Voltage Unbalance for. Wilson, Gardner.......426–37; disc. 437
Time Scale and Gain on Criteria for Servomechanism Performance, The Influence of. Graham, Lathrop.......153–7; disc. 157
Tinning, Trends in "Automation"—Electrolytic. Gravenstreter, Layton.......97–100; disc. 100
Tools, Recent Developments in the Theory and Design of Electric Spark Machine. Williams, Woodford, Smith.......83–6; disc. 87
Toronto Subway, The. Inglis.......416–20; disc. 420
Traction Motor for Diesel-Electric Locomotives, A World-Wide Standard. Baldwin.......144–7
Transfer Functions, Self-Oscillation Method for Measuring. Clegg, Harris.......169–70
Transformation Theory for Sampled-Data Systems, Additions to z-. Lago.......403–08
Transformerless Rectified Higher Voltage D-C Aircraft Electric System, Study of. Dallas, Reising.......253–7; disc. 257
Transient Analysis of Sampled-Data Control Systems. Johnson, Lindorff.......147–53
Transient Characteristics of Aircraft A-C Generators. Holloway.......187–90
Transient Performance of Servomechanisms with Derivative and Integral Control, The. Lathrop, Graham.......10–17
Transit, An Electric System for Monorail Rapid. Anson, Kimball.......225–6; disc. 226
Transit Equipment, High-Speed Rapid. Vouch.......267–70; disc. 270

Transit System, Rebuilding San Francisco's. ReQua.............171–6
Transparent Conductive Films, Aircraft Windshields Heated by Means of. Ward..............408–16
(Transportation, Air) A High-Power Servo Analyzer. Gorrill, Walley..........................338–43
(Transportation, Air) A Weight Analysis of Modern Aircraft Electric Systems. Garbarino, Hawkes, Granath..........................463–8; disc. 468
(Transportation, Air) Aircraft Tachometer Indicator—An Analysis of Design Factors Affecting Starting Performance. Akeley..................398–403
(Transportation, Air) Aircraft Windshields Heated by Means of Transparent Conductive Films. Ward...408–16
(Transportation, Air) Carbon-Pile Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation. Richards....................357–63
(Transportation, Air) Characteristics of Aircraft A-C Generators. Stratton, Matsch.............165–9
(Transportation, Air) Comments on Aircraft Switch Testing. Stuelpnagel, Dallas.............301–05
(Transportation, Air) Development of a Large Ratio-Release Circuit-Breaker Mechanism. Scarcelli, Steiner.......................512–17; disc. 517
(Transportation, Air) Development of an Air-Borne Stabilized Camera Mount. Miller, Alexander...184–6
(Transportation, Air) Economic Factors for Aircraft Electric Power Systems. Bergslien, Stratton, Finison.......................270–8; disc. 278
(Transportation, Air) Environmental Testing of Small Electronic Components at High Sound Intensities. Mintz, Levine................503–09; disc. 509
(Transportation, Air) Evaluation of Designs for Intermittently Heated Surfaces. Dahm, Holloway...198–205
(Transportation, Air) Experimental Determination of 400-Cycle Impedance of Wire in Aircraft Power Distribution Circuits. Andrew...469–77; disc. 477
(Transportation, Air) High Altitude Aircraft Inverters. Franklin..............................363–72
(Transportation, Air) Overvoltage Detection in Single and Multigenerator Aircraft A-C Systems. Tucker, Trbovich.......................420–4; disc. 424
(Transportation, Air) Performance of a Constant-Speed Drive. Giloy..........................179–84
(Transportation, Air) Reliability Improvement—A Plan to Achieve and Measure Reliability of Aircraft Equipment. Luckman.................378–83
(Transportation, Air) Sensitivity Requirements of Reactive Load Division Circuits in Aircraft Electric Systems. Sherrard.............279–86; disc. 286
(Transportation, Air) Thermal Developments in Aircraft Generators. Martin...........493–502; disc. 502
(Transportation, Air) Transient Characteristics of Aircraft A-C Generators. Holloway..........187–90
(Transportation, Land) A World-Wide Standard Traction Motor for Diesel-Electric Locomotives. Baldwin...................................144–7
(Transportation, Land) Diesel-Electric Locomotive Wheel Slipping—Causes, Effects, and Methods of Control. Fort..................478–80; disc. 480
(Transportation, Land) Fundamentals of Flashing of Diesel-Electric Motors and Generators. Atwell.......................................70–6; disc. 76
(Transportation, Land) Gearing for Diesel-Electric Locomotives. Bevan...........510–12; disc. 512
(Transportation, Land) Maintenance Testing of Insulation Resistance on Diesel-Electric Locomotives. Kelley..........................452–4; disc. 454
(Transportation, Land) Rectifier Motive Power—Inductive Co-ordination Considerations. King, Gordon, Hibbard.......................107–14; disc. 114
(Transportation, Land) Some Application Phases of the Ignitron Rectifier Locomotives on the Pennsylvania Railroad. Brown..............128–33; disc. 133
(Transportation, Land) Static Excitation Control for Diesel-Electric Locomotives. McElhenny, Smith...158–60
(Transportation, Land) The Toronto Subway. Inglis.............................416–20; disc. 420
Trends in "Automation"—Electrolytic Tinning. Gravenstreter, Layton...........97–100; disc. 100
208-Volt Service in Commercial Buildings, The Relative Feasibility of 460-Volt or. Barnett, Zimmerman, Lokay.......................306–13; disc. 337
Two-Phase Servomotor, Evaluating the Effect of Nonlinearity in a. Stein, Thaler....518–21; disc. 521

U

Unbalance for 3-Phase Synchronous Systems, Calculations on Voltage. Wilson, Gardner...426–37; disc. 437
(United States Steel) A-C Power System at the Fairless Works. Watkins, Derr, Fountain, Squires..342–52
Unity, A Graphical Procedure for Determining the Gain of a Servomechanism for a Specified Maximum Modulus Less Than. Higgins...101–03; disc. 103
Utility Power Systems, Problems Relating to Interconnections of Large Pulp and Paper Mills with Large. Rose, Springer.................205–14; disc. 214
Utilization Equipment, Service Voltage Spread and Its Effect on. Barnett, Lawrence....320–5; disc. 336

V

Voltage, A Self-Excited Induction Generator with Regulated. McConnell.............288–94; disc. 294
Voltage for Metropolitan Areas, Symposium on Higher Distribution...................306–35; disc. 335
Voltage Ripple on D-C Generators, The Nature of. Collamore.............................390–3
Voltage Spread and Its Effect on Utilization Equipment, Service. Barnett, Lawrence......320–5; disc. 336
Voltage Unbalance for 3-Phase Synchronous Systems, Calculations on. Wilson, Gardner...426–37; disc. 437

W

Weight Analysis of Modern Aircraft Electric Systems, A. Garbarino, Hawkes, Granath.....463–8; disc. 468
Wheel Slipping—Causes, Effects, and Methods of Control, Diesel-Electric Locomotive. Fort...478–80; disc. 480
Windshields Heated by Means of Transparent Conductive Films, Aircraft. Ward..............408–16
Wire in Aircraft Power Distribution Circuits, Experimental Determination of 400-Cycle Impedance of. Andrew.......................469–77; disc. 477
World-Wide Standard Traction Motor for Diesel-Electric Locomotives, A. Baldwin................144–7

Z

z-Transformation Theory for Sampled-Data Systems, Additions to. Lago.....................403–08

2. Author Index

Akeley, L. T. Aircraft Tachometer Indicator—An Analysis of Design Factors Affecting Starting Performance. (54–358)..................398–403
Alden, Carroll R. Disc..........................87
Alexander, A. J.; J. H. Miller. Development of an Air-Borne Stabilized Camera Mount. (54–310)..184–6
Alexander, F. Disc..........................214
Andrew, J. D. Experimental Determination of 400-Cycle Impedance of Wire in Aircraft Power Distribution Circuits. (54–374)..............469–77
Anger, E. G.; D. L. Pettit. Accurate Control of Relative Speed and Cut in a Continuous Process Line. (54–387)......................485–90; disc. 492
Anson, E. H.; R. L. Kimball. An Electric System for Monorail Rapid Transit. (54–251)...225–6; disc. 226
Armstrong, Thomas W. Disc....................278
Asbury, C. E.; O. A. Lentz. Disc...............335
Atwell, C. A. Fundamentals of Flashing of Diesel-Electric Motors and Generators. (54–45)..70–6; disc. 76

B

Baldwin, M. J. A World Wide Standard Traction Motor for Disel-Electric Locomotives. (54–189)...144–7
Barker, R. D. Disc..........................246
Barnett, H. G.; R. F. Lawrence. Service Voltage Spread and Its Effect on Utilization Equipment. (S-66)......................320–5; disc. 336
Barnett, H. G.; R. A. Zimmerman, H. E. Lokay. The Relative Feasibility of 460-Volt or 208-Volt Service in Commercial Buildings. (S-66).............................306–13; disc. 337
Baumann, G. A. Disc..........................490
Beck, C. D. Disc..........................491
Beckett, J. C. Air Ionization as an Environment Factor. (54–225)..............................161–5
Beeman, Donald; H. D. Kurt. Progress in Power System Engineering for Commercial Buildings. (S-66)..........................313–20; disc. 335
Bergen, Arthur R.; J. R. Ragazzini. Sampled-Data Processing Techniques for Feedback Control Systems. (54–282)................236–44; disc. 246
Bergslien, R. M.; L. J. Stratton, H. J. Finison. Economic Factors for Aircraft Electric Power Systems. (54–306)......................270–8; disc. 278
Bevan, G. T. Gearing for Diesel-Electric Locomotives. (54–382)..............................510–12
Bisson, A. J.; E. A. Rochau. Iron Conduit Impedance Effects in Ground Circuit Systems. (54–36)...104–06
Boggess, P. F. Disc........................263, 286
Bogner, Irving; L. F. Kazda. An Investigation of the Switching Criteria for Higher Order Contactor Servomechanisms. (54–44).....118–26; disc. 126
Brereton, D. S.; H. J. Donnelly. 480 Wye/277-Volt Power System in Telephone Building at Menands, N.Y. (S-66)..................325–35; disc. 338
Brieger, L. Disc..........................231
Brown, F. D. Some Application Phases of the Ignitron Rectifier Locomotives on the Pennsylvania Railroad. (54–116)................128–33; disc. 135
Brown, H. F. Disc..............76, 114, 134, 461, 512
Brown, H. F.; R. L. Kimball. A Reappraisal of the Economics of Railway Electrification: How, When, and Where Can It Compete with the Diesel-Electric Locomotive? (54–29)............35–45; disc. 48
Bruma, M. S. Disc..........................87
Bucci, N. W., Jr. Disc..........................279
Buck, A. M. Disc..........................215
Bullard, W. R. Disc..........................337
Burnett, J. R.; P. E. Kendall. Linear Compensation of Saturating Servomechanisms. (54–125).....6–10
Byloff, R. C.; E. D. Lytle, J. B. Caldwell. Disc.....294

C

Caldwell, J. B.; E. D. Lytle, R. C. Byloff. Disc....294
Caldwell, S. C.; L. E. Jensen. A Method of Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems. (54–307)..............263–7; disc. 267
Calvert, J. F.; D. J. Ford. The Application of Short-Time Memory Devices to Compensator Design. (54–124)........................88–93; disc. 93
Cataldo, J. B.; N. Shackman. Current Distribution in Paired-Phase Bus Bars Under Unbalanced Load Conditions. (54–329)..........438–42; disc. 443
Cham, E. J.; R. J. Radus. Disc.................96
Cham, E. J.; G. M. Zins. Considerations in Applying Rectifiers as a Power Supply for Hot Strip Mills. (54–107)........................65–9; disc. 70
Chow, C. K. Contactor Servomechanisms Employing Sampled Data. (54–123).........51–62; disc. 64
Clegg, John C., L. D. Harris. Self-Oscillation Method for Measuring Transfer Functions. (54–284)..169–70
Coleman, H. C. Disc..........................138
Collamore, F. N. The Nature of Voltage Ripple on D-C Generators. (54–420)...................390–3
Conlon, William J. Disc..........................80
Craig, D. Norman; W. J. Hamer. Some Aspects of the Charge and Discharge Processes in Lead Acid Storage Batteries. (54–180)..................22–34
Curry, H. H. Disc..........................139

D

Dahm, Thomas M.; R. A. Holloway. Evaluation of Designs for Intermittently Heated Surfaces (54–308)..198–205
Dale, E. E. Disc..........................214
Dallas, J. P.; C. A. Reising, Jr. Study of Transformerless Rectified Higher Voltage D-C Aircraft Electric System. (54–309)..............253–7; disc. 257
Dallas, J. P.; T. R. Stuelpnagel. Comments on Aircraft Switch Testing. (54–352)................301–05
DeBarber, J. P. A Magnetic Tape Memory for D-C Positional Servomechanisms. (54–520).....372–4
Derr, W. A.; M. A. Hyde. Remote Operation of Pipe-Line Pumping Stations. (54–221).........190–8
Derr, W. A.; S. S. Watkins, L. L. Fountain, R. B. Squires. A-C Power System at the Fairless Works. (54–348)..............................343–52
Donnelly, H. J.; D. S. Brereton. 480 Wye/277-Volt Power System in Telephone Building at Menands, N.Y. (S-66)..................325–35; disc. 338
Downing, E. A. Magnetic Amplifiers in Metering Direct Current on Electrolytic Cell Lines (54–178)..............................93–6; disc. 96

E

Evans, S. O. Disc..........................356
Exner, D. W. Disc..........................468, 517
Exner, D. W.; B. V. Hoard. Disc..............477

F

Ferri, L. F. Disc..........................80
Fifer, W. H. Disc..........................138
Finison, H. J.; R. M. Bergslien, L. J. Stratton. Economic Factors for Aircraft Electric Power Systems. (54–306)......................270–8; disc. 278

Fisher, Lawrence E. Disc. 442

Ford, D. J.; J. F. Calvert. The Application of Short-Time Money Devices to Compensator Design (54-124) ....88–93; disc. 93

Fort, R. I. Diesel-Electric Locomotive Wheel Slipping—Causes, Effects, and Methods of Control. (54-350) ....478–80; disc. 480

Fountain, L. L.; . S. S. Watkins, W. A. Derr, R. B. Squires. A-C Power System at the Fairless Works. (54–348) 343–52

Franklin, P. W. High Altitude Aircraft Inverters. (54-415) 363–72

Freeman, G.; M. Pella. Disc. 509

Frick, C. W. Disc. 116

Friedman, Daniel. Disc. 443

Fuchs, Abraham M. Disc. 93, 483

G

Garbarino, H. L.; A. K. Hawkes, J. A. Granath. A Weight Analysis of Modern Aircraft Electric Systems. (54–357) ....463–8; disc. 468

Gardner, W. K.; B. J. Wilson. Calculations on Voltage Unbalance for 3-Phase Synchronous Systems. (54-359) ....426–37; disc. 437

Garry, F. T. Disc. 115, 133

Gehmlich, Dietrich K.; M. E. Van Valkenburg. Measurement of Some Nonlinearities in Servomechanisms. (54–285) 232–5

Gibson, G. W. Disc. 80

Gienger, J. A. Disc. 231

Giloy, E. W. Performance of a Constant-Speed Drive. (54–217) 179–84

Giloy, E. W.; M. J. Powell. Considerations Applicable to Automatic Paralleling of Aircraft A-C Generators. (54–312) ....258–63; disc. 263

Gordon, K. H.; E. B. King, L. J. Hibbard. Rectifier Motive Power—Inductive Co-ordination Considerations. (54–110) ....107–14; disc. 117

Gorrill, W. Sterling; O. C. Walley. A High-Power Servo Analyzer. (54–419) 338–43

Graham, Dunstan; R. C. Lathrop. The Influence of Time Scale and Gain on Criteria for Servomechanism Performance. (54–197) .....153–7; disc. 158

Graham, Dunstan; R. C. Lathrop. The Transient Performance of Servomechanisms with Derivative and Integral Control. (54–126) 10–17

Granath, J. A.; H. L. Garbarino, A. K. Hawkes. A Weight Analysis of Modern Aircraft Electric Systems. (54–357) ....463–8; disc. 468

Gravenstreter, Paul R.; R. E. Layton. Trends in "Automation"—Electrolytic Tinning. (54–108). ....97–100; disc. 100

Graybill, H. W. A New High-Current Switch for Electrolytic and Electrothermal Applications. (54–176) 1–6

Gross, Eric T. B. Disc. 478

H

Haas, Vinton B., Jr. Disc. 235

Hamer, Walter J.; D. N. Craig. Some Aspects of the Charge and Discharge Processes in Lead Acid Storage Batteries. (54–180) 22–34

Hamilton, W. S. H. Disc. 45, 455

Harris, L. Dale; J. C. Clegg. Self-Oscillation Method for Measuring Transfer Functions. (54–284) 169–70

Hawkes, A. K.; H. L. Garbarino, L. A. Granath. A Weight Analysis of Modern Aircraft Electric Systems. (54–357) ....463–8; disc. 468

Heiss, F. D. Disc. 117

Hibbard, L. J.; E. B. King, K. H. Gordon. Rectifier Motive Power—Inductive Co-ordination Considerations. (54–110) ....107–14; disc. 117

Higgins, Thomas J. A Graphical Procedure for Determining the Gain of a Servomechanism for a Specified Maximum Modulus Less Than Unity. (54–127) ....101–03; disc. 103

Higgins, Thomas J. Disc. 442

Hines, C. M. Disc. 420

Hoard, Bert V. Disc. 437

Hoard, Bert V.; D. W. Exner. Disc. 477

Holden, W. H. T. Disc. 226, 270

Holloway, Raymond A.; T. M. Dahm. Evaluation of Designs for Intermittently Heated Surfaces. (54-308) 198–205

Holloway, V. C. Transient Characteristics of Aircraft A-C Generators. (54–305) 187–90

Hovious, R. L. Jitter in Instrument Servos. (54–521) 393–8

Hoxie, E. A. Some Discharge Characteristics of Lead Acid Batteries. (54–177) 17–22

Huldschiner, Dr. G. Disc. 46

Hunkins, H. D. Disc. 215

Hyde, M. A.; W. A. Derr. Remote Operation of Pipe-Line Pumping Stations. (54–221) 190–8

I

Inglis, J. G. The Toronto Subway. (54–343) ....416–20; disc. 420

J

Jackson, Albert S. An Extension of the Root Locus Method to Obtain Closed-Loop Frequency Response of Feedback Control Systems. (54–281) 176–9

Jenkins, C. F. Disc. 76

Jensen, L. E.; S. C. Caldwell. A Method of Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems. (54–307) ....263–7; disc. 267

Johnson, G. W.; D. P. Lindorff. Transient Analysis of Sampled-Data Control Systems. (54–196) 147–53

Jury, Eliahu I. Disc. 252

K

Kalman, Rudolf E. Phase-Plane Analysis of Automatic Control Systems with Nonlinear Gain Elements. (54–319) 383–90

Kalman, R. E. Disc. 245

Kaufmann, R. H. Some Fundamentals of Equipment-Grounding Circuit Design. (54–244) ....227–31; disc. 232

Kaufmann, R. H. Disc. 81, 107

Kazda, Louis F.; I. Bogner. An Investigation of the Switching Criteria for Higher Order Contactor Servomechanisms. (54–44) .....118–26; disc. 126

Kelley, W. E. Maintenance Testing of Insulation Resistance on Diesel-Electric Locomotives. (54–339) ....452–4; disc. 454

Kendall, P. E.; J. R. Burnett. Linear Compensation of Saturating Servomechanisms. (54–125) 6–10

Kimball, R. L.; E. H. Anson. An Electric System for Monorail Rapid Transit. (54–251) ....225–6; disc. 226

Kimball, R. L.; H. F. Brown. A Reappraisal of the Economics of Railway Electrification: How, When, and Where Can It Compete with the Diesel-Electric Locomotive? (54–29) ....35–45; disc. 48

Kincaid, N. L.; R. H. Wright. Fairless Works—Electric Equipment for Slabbing Mill and Blooming Mill. (54–181) ....141–3; disc. 143

King, E. B.; K. H. Gordon, L. J. Hibbard. Rectifier MotivePower—Inductive Co-ordination Considerations. (54–110) ....107–14; disc. 117

Kirsch, William A. Disc. 480

Kramer, C. D. Induction Heating Steel with 60 Cycles. (54–465) ....353–5; disc. 356

Kullrich, W. Disc. 216

Kurt, H. D.; D. Beeman. Progress in Power System Engineering for Commercial Buildings. (S-66) ....313–20; disc. 335

L

Lago, Gladwyn V. Additions to z-Transformation Theory for Sampled-Data Systems. (54–523) 403–08

Lago, Gladwyn V.; J. G. Truxall. The Design of Sampled-Data Feedback Systems. (54–283) ....247–53; disc. 253

Lange, A. Disc. 126

Larson, H. E. Disc. 69, 143

Larson, Lee R. Disc. 468

Larson, R. A. Disc. 294

Lathrop, Richard C.; D. Graham. The Influence of Time Scale and Gain on Criteria for Servomechanism Performance. (54–197) .....153–7; disc. 158

Lathrop, Richard C.; D. Graham. The Transient Performance of Servomechanisms with Derivative and Integral Control. (54–126) 10–17

Lawrence, Robert F. Disc. 335

Lawrence, R. F.; H. G. Barnett. Service Voltage Spread and Its Effect on Utilization Equipment. (S-66) ....320–5; disc. 336

Layton, Robert E.; P. R. Gravenstreter. Trends in "Automation"—Electrolytic Tinning. (54–108). ....97–100; disc. 100

Lebens, John C. Co-ordinated Fuse Protection for Low-Voltage Distribution Systems in Industrial Plants. (54–9) ....77–80; disc. 81

Lee, Reuben; G. R. Mohr. Magnetic Amplifier Control of Radio-Frequency Generators. (54–377) 374–8

Legg, L. Elton. Operating Experience in Diesel-Electric Locomotives Results in Design Changes. (54–344) ....461–3; disc. 463

Lentz, O. A.; C. E. Asbury. Disc. 335

Levine, M. B.; F. Mintz. Environmental Testing of Small Electronic Components at High Sound Intensities. (54–422) ....503–09; disc. 509

Lindorff, D. P.; G. W. Johnson. Transient Analysis of Sampled-Data Control Systems. (54–196) . .147–53

Lokay, H. E.; H. G. Barnett, R. A. Zimmerman. The Relative Feasibility of 460-Volt or 208-Volt Service in Commercial Buildings. (S-66) .306–13; disc. 337

Lokay, H. E.; R. A. Zimmerman. Disc. 337

Luckman, J. E. Reliability Improvement—A Plan to Achieve and Measure Reliability of Aircraft Equipment. (54–353) 378–83

Lytle, E. D.; J. B. Caldwell, R. C. Byloff. Disc. 294

M

MacLeod, D. R. Disc. 115, 134

Magat, M. Disc. 87

Martin, Cecil G. Thermal Developments in Aircraft Generators. (54–414) ....493–502; disc. 502

Martin, Thomas J. Disc. 517

Matsch, L. W.; L. J. Stratton. Characteristics of Aircraft A-C Generators. (52–319) 165–9

Matthysse, I. Disc. 267

McArthur, B. E. Disc. 356

McConnell, H. M. A Self-Excited Induction Generator with Regulated Voltage. (54–317) ....288–94; disc. 295

McElhenny, S. W.; R. M. Smith. Static Excitation Control for Diesel-Electric Locomotives. (54–190) 158–60

Miller, J. H.; A. J. Alexander. Development of an Air-Borne Stabilized Camera Mount. (54–310) .184–6

Mintz, F.; M. B. Levine. Environmental Testing of Small Electronic Components at High Sound Intensities. (54–422) ....503–09; disc. 509

Mohr, G. R.; R. Lee. Magnetic Amplifier Control of Radio-Frequency Generators. (54–377) 374–8

Monfort, S. A. Disc. 294

Murphy, E. K. Disc. 215

N

Naumann, Erwin O. A. Correlation of Rating Data for Rotating Electric Machinery. (54–355) ....443–51; disc. 451

O

Ogden, H. S. Disc. 116, 134

Oldenburger, Rufus. Disc. 126

Oscarson, G. L. Application of Large A-C Motors and Controls in Refineries and Pipe Lines. (54–526) 296–301

Owen, T. B. Disc. 424

P

Pansini, A. J. Disc. 337

Pella, M.; G. Freeman. Disc. 509

Perkinson, T. F. Disc. 46

Petree, Frank. Disc. 103

Pettit, D. L.; E. G. Anger. Accurate Control of Relative Speed and Cut in a Continuous Process Line. (54–387) ....485–90; disc. 492

Pfeiffer, Paul E. The Maximum Response Ratio of Linear Systems. (54–371) ....480–3; disc. 483

Pfeiffer, P. E. Disc. 157

Powell, M. J.; E. W. Giloy. Considerations Applicable to Automatic Paralleling of Aircraft A-C Generators. (54–312) ....258–63; disc. 263

Prince, Luther T., Jr. A Generalized Method for Determining the Closed-Loop Frequency Response of Nonlinear Systems. (54–280) . . .217–24; disc. 224

R

Radus, R. J.; E. J. Cham. Disc. 96

Ragazzini, John B.; A. R. Bergen. Sampled-Data Processing Techniques for Feedback Control Systems. (54–282) ....236–44; disc. 246

Reising, C. A., Jr.; J. P. Dallas. Study of Transformerless Rectified Higher Voltage D-C Aircraft Electric System. (54–309) ....253–7; disc. 257

ReQua, F. L. Rebuilding San Francisco's Transit System. (54-253) 171–6

Richards, H. H. C. Carbon-Pile Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation. (54–356) 357–63

Rochau, E. A.; A. J. Bisson. Iron Conduit Impedance Effects in Ground Circuit Systems. (54–36) 104–06

Rose, H. A.; H. E. Springer. Problems Relating to Interconnections of Large Pulp and Paper Mills with Large Utility Power Systems. (54–243) ....205–14; disc. 216

Russell, Frederick A. Disc. 62

Rutan, E. J. Disc. 107

S

Salzer, John M. Disc. 244, 252
Saunders, R. M. Disc. 224
Sauter, D. M. Disc. 336
Scarcelli, J. A.; R. Steiner. Development of a Large Ratio-Release Circuit-Breaker Mechanism. (54–423) .... 512–17; disc. 518
Schach, M.; D. H. Scott. Disc. 278
Schaelchlin, W. Disc. 138
Schmidt, A., Jr. Disc. 69
Scott, D. H. Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled Aircraft Generators. (54–346) .... 455–60; disc. 461
Scott, D. H.; M. Schach. Disc. 278
Shackman, N.; J. B. Cataldo. Current Distribution in Paired-Phase Bus Bars Under Unbalanced Load Conditions. (54–329) .... 438–42; disc. 443
Sherrard, Emile S. Sensitivity Requirements of Reactive Load Division Circuits in Aircraft Electric Systems. (54–311) .... 279–86; disc. 286
Smith, Richard E.; E. M. Williams, J. B. Woodford, Jr. Recent Developments in the Theory and Design of Electric Spark Machine Tools. (54–174) .... 83–6; disc, 88
Smith, R. M.; S. W. McElhenny. Static Excitation Control for Diesel-Electric Locomotives. (54–190) 158–60
Springer, H. E.; H. A. Rose. Problems Relating to Interconnections of Large Pulp and Paper Mills with Large Utility Power Systems. (54–243) .... 205–14; disc. 216
Squires, R. B.; S. S. Watkins, W. A. Derr, L. L. Fountain. A-C Power System at the Fairless Works. (54–348) 343–52
Stein, W. A.; G. J. Thaler. Evaluating the Effect of Nonlinearity in a 2-Phase Servomotor. (54–522) .... 518–21; disc. 521
Steiner, Rudolf; J. A. Scarcelli. Development of a Large Ratio-Release Circuit-Breaker Mechanism. (54–423) .... 512–17; disc. 518
Stineman, Russell W. Disc. 263, 286, 424
Stone, Kenneth A. Disc. 521
Stotz, J. K. Disc. 76
Stratton, L. J.; R. M. Bergslien, H. J. Finison. Economic Factors for Aircraft Electric Power Systems. (54–306) .... 270–8; disc. 278
Stratton, L. J.; L. W. Matsch. Characteristics of Air craft A-C Generators. (52–319) 165–9
Stuelpnagel, T. R.; J. P. Dallas. Comments on Aircraft Switch Testing. (54–352) 301–05
Summers, O. R. Disc. 356

T

Thaler, G. J.; W. A. Stein. Evaluating the Effect of Nonlinearity in a 2-Phase Servomotor. (54–522) .... 518–21; disc. 521
Thomas, C. C. Disc. 100
Thompson, Fred. Disc. 215
Tomalin, Paul G. Disc. 140
Trbovich, M.; W. M. Tucker. Overvoltage Detection in Single and Multigenerator Aircraft A-C Systems. (54–327) .... 420–4; disc. 425
Truxall, John G.; G. V. Lago. The Design of Sampled-Data Feeback Systems. (54–283) .... 247–53; disc. 253
Tucker, W. M.; M. Trbovich. Overvoltage Detection in Single and Multigenerator Aircraft A-C Systems. (54–327) .... 420–4; disc. 425

V

VanValkenburg, M. E.; D. K. Gehmlich. Measurement of Some Nonlinearities in Servomechanisms. (54–285) 232–5
Vouch, S. J. High-Speed Rapid Transit Equipment. (54–252) .... 267–70; disc. 270
Vredenburgh, E. J. Disc. 103

W

Walley, Omar C.; W. S. Gorrill. A High-Power Servo Analyzer. (54–419) 338–43
Ward, John W. Aircraft Windshields Heated by Means of Transparent Conductive Films. (54–525) 408–16
Wasmund, James A. Series- Versus Parallel-Connected Generators for Multiple-Engine D-C Diesel-Electric Ship-Propulsion Systems. (54–144) .... 135–8; disc. 140
Watkins, S. S.; W. A. Derr, L. L. Fountain, R. B. Squires. A-C Power System at the Fairless Works. (54–348) 343–52
Webb, R. L. Disc. 45
Whitaker, Thomas N. Disc. 484
Whittaker, C. C. Disc. 117
Wilde, O. A. Disc. 139
Williams, E. M.; J. B. Woodford, Jr., R. E. Smith. Recent Developments in the Theory and Design of Electric Spark Machine Tools. (54–174) .... 83–6; disc. 88
Wilson, B. J. Disc. 257
Wilson, B. J.; W. K. Gardner. Calculations on Voltage Unbalance for 3-Phase Synchronous Systems. (54–359) .... 426–37; disc. 437
Withington, Sidney. Disc. 47
Woodford, J. B., Jr.; E. M. Williams, R. E. Smith. Recent Developments in the Theory and Design of Electric Spark Machine Tools. (54–174) .... 83–6; disc. 88
Wright, R. H.; N. L. Kincaid. Fairless Works—Electric Equipment for Slabbing Mill and Blooming Mill. (54–181) .... 141–3; disc. 143

Y

Yereance, Robert E. Disc. 460, 502

Z

Zimmerman, R. A.; H. G. Barnett, H. E. Lokay. The Relative Feasibility of 460-Volt or 208-Volt Service in Commercial Buildings. (S-66) .. 306–13; disc. 337
Zimmerman, R. A.; H. E. Lokay. Disc. 337
Zins, G. M.; E. J. Cham. Considerations in Applying Rectifiers as a Power Supply for Hot Strip Mills. (54–107) .... 65–9; disc. 70

A New High-Current Switch for Electrolytic and Electrothermal Applications

H. W. GRAYBILL
MEMBER AIEE

THE marked increase in electrolytic cell-operating currents in the past decade has emphasized the need for a simple, dependable, and inexpensive high-current switch for use as a cell short-circuiting switch. In the electrothermal field, the magnitude of furnace currents has also been increasing as it has become evident that higher furnace currents lead to better operating economies. Earlier types of switches become cumbersome and expensive when built in ratings of higher than about 6,000 amperes. This paper traces the development of a new type of switch and illustrates how the new switch is applicable to both a-c and d-c low-voltage circuits.

Requirements of a Cell Short-Circuiting Switch

Electrolytic cells, in general, operate at a very low voltage and a relatively high current. In certain metal-refining operations, the voltage drop across a cell is as low as a fraction of a volt, but in most electrolytic processes the cell voltage runs from 3 to 5 volts. In extreme cases, voltages as high as 20 or 25 volts across a single cell have been reported. Most commercial electrolytic processes operate at a minimum of 3,000 or 4,000 amperes per cell, and currents of 20,000 to 36,000 amperes are now commonplace in chlorine cells. Processes using fused electrolytes such as the reduction of aluminum, magnesium, and sodium, employ currents ranging as high as 100,000 amperes. In view of the high currents and low voltages involved, it is most economical to connect a large number of cells in series, so as to obtain a total voltage for the cell line of from 200 to 700 volts.

Because of the series connection, it is necessary to short-circuit individual cells whenever a cell is required to be removed from service. This may be for routine maintenance and repairs, or it may be during the tapping operation, or, in other cases, it may be necessary to short-circuit the cell very quickly after failure of auxiliary apparatus, to prevent unwanted or disastrous results. The requirement for frequency of operation, therefore, varies from daily for some types of cells to semi-annually or annually for other types of cells. In every case, the short-circuiting switch is required to carry the full rated current of the cell line for as long as several days and to interrupt that current at the cell-operating voltage each time the switch is opened. If the switch is to be operated frequently, it should obviously be so designed and constructed that it can be quickly and easily operated by one man, to reduce operating cost to a minimum and to minimize lost production time. It should be so arranged that it is impossible to leave a part of the switch in the open position while another part is in the closed position, carrying the full cell current. It should also be independent of the human element or the operator's judgment. For example, the operator should not have to guess how tight the bolts or screws should be tightened to obtain the proper contact pressure. To inhibit contact corrosion, contacts must be faced with a noble metal such as silver, and contacts should make and break with a wiping motion to rub away deposits of dust and foreign material. Above all, the device must be simple and reliable in its operation, and require a minimum of maintenance.

The service conditions under which the cell switch has to operate are some of the worst encountered by any electrical device. In various installations, salts, acids, and alkalis are spilled over the switch, and chlorine, sulphurous compounds, and other corrosive gases are present in the atmosphere. Many operate in a high ambient temperature, connected to bus which is running at a temperature well above the 30 degrees centigrade rise which is recognized as good operating practice in standards of the AIEE and the National Electrical Manufacturers Association. On switches that are operated daily, the interruption of cell current each time the switch opens causes pitting of the contacts which leads to a maintenance problem.

Early Types of Short-Circuiting Devices

Perhaps the earliest cell short-circuiting device in use was the simple disconnecting link, and such links are still in use today. Two or more links are often used in parallel, as required to carry the rated current, and are usually bolted in place at both the hinge and break ends. The links are rotated about the bolt at one end. After they have been swung into the closed position, the bolts are tightened to give high-pressure contact at both the hinge and break ends. Sometimes two or three bolts may be used at each end of the rotating link. This device has the following disadvantages:

1. Length of time required to operate.
2. Where two or more links are used in parallel, it may be possible for the operator to leave some links open and others closed.
3. The last link or links to open usually draw a heavy spark or arc which may make it necessary to file or grind the parting edges of the link before the device can be closed again.
4. Where heavy cross sections of copper are required for the high currents, alignment becomes a very definite problem. A poorly aligned switch may overheat to the extent that, in the extreme case, the copper blades may actually melt.

Another early short-circuiting device was the wedge. A wedge-shaped piece of conducting material, usually copper, is driven into and out of engagement with a wedge-shaped gap in the bus, terminals of which are connected to each polarity of the cell or pot. Like the bolted link, the wedge has the advantage of low first cost and, like the bolted link,

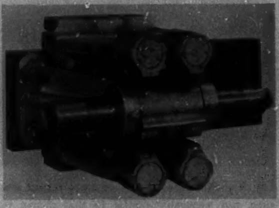


Fig. 1. Early Canadian cell short-circuiting switch

Paper 54-176, recommended by the AIEE Chemical, Electrochemical and Electrothermal Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing November 30, 1953.

H. W. Graybill is with the I-T-E Circuit Breaker Company, Greensburg, Pa.

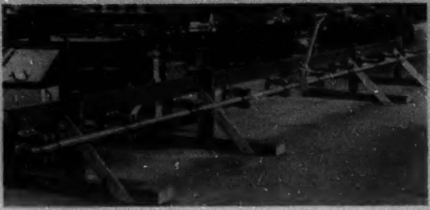


Fig. 2. Shop setup of 25,000-ampere cell short-circuiting switch using the switch units shown in Fig. 1

Fig. 4. Shop setup of 10,000-ampere cell short-circuiting switch using switch units shown in Fig. 3

it is still widely used today, particularly on aluminum pots. Unfortunately, since the wedge is a solid piece of material, it is not possible to have contact between the wedge and the bus at more than one or two points on each side. For this reason, the wedge is suitable for carrying the full cell current for a limited time only. For longer periods, bolted links or plates must be added in parallel with the wedge.

An entirely new type of cell short-circuiting switch was introduced in Canada about 1947. A number of these switches were installed in a chlorine plant in Quebec about that time and have given very satisfactory service. This switch, which is shown in Fig. 1, was an adaption of the familiar rotating flat-contact (beaver-tail) switch used in outdoor gang-operated switches. A copper casting of a generally elliptical shape was mounted on a rotating steel shaft in such a manner as to rotate into and out of contact with stationary jaws. The cast-copper jaws were mounted on an insulating base, as were the bearings. Contact pressure was supplied by heavy U-shaped springs which were insulated at either end so that the top contacts were electrically insulated from the bottom contacts. The switch unit shown was rated 2,500 amperes, and up to ten of these units could be operated from a single handle simultaneously, as shown in Fig. 2. In service, all of these units are connected in parallel, electrically. We believe that this switch was the first standard-production model designed in units that could be placed in any physical arrangement and connected in parallel to give any desired current rating, with all of the units operated simultaneously.

An Italian design, introduced about the same time or perhaps earlier, consisted essentially of a 1/2-inch thick by 6-inch square silver-plated copper bar which was inserted, flatwise, between two 1/2-inch by 6-inch bus bars, one of which was connected to the cell anode and the other to the cathode. Contact pressure was supplied by a 1-inch bolt and a heavy Belleville washer. A number of these switch units were also connected to operate in parallel and were opened and closed by push-pull rods operated from a crank-and-worm gear arrangement. The chief disadvantage of this switch was the large contact area and the resulting low-unit contact pressure.

Fig. 3 shows a switch which was introduced in this country in 1949. This switch was the same in principle as that shown in Fig. 1, but hard-drawn copper stationary contact fingers were substituted for the cast contacts used in the Canadian design. The fingers and shaft bearings were again mounted on an insulating base. The U-shaped springs were insulated from the copper contact fingers by means of the molded phenolic blocks which can be seen in the foreground. The $1^1/_2$-inch diameter shaft was rotated 60 degrees from the closed position to the open position. Shafts of adjacent switch units were connected with lengths of tubing, jig-drilled to be pinned to each shaft. On the end switch unit shown in this figure, an auxiliary switch was added to give a remote position indication. The switch was enclosed in a cast drip-proof housing. Fig. 4 shows a shop setup of the 10,000-ampere, 4-unit switch with operating handle. The operating handle was arranged to be removable in the open position, which is shown in Fig. 4. This was necessary because in this position, the operating handle blocks the aisle between the cells. In the closed position, this handle stands vertically and is retained in place so that the switch can be closed quickly when and if required. The switch handle also gives an indication of switch position which can be seen from a distance.

Several hundred of these switches are installed in current ratings ranging from 10,000 to 30,000 amperes, in California,

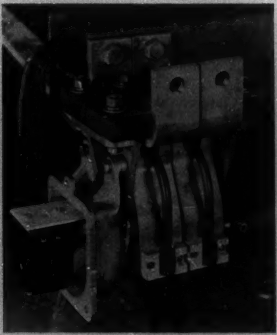


Fig. 3. American switch of more recent design, but generally similar to that shown in Fig. 1. Auxiliary switch attachment also shown

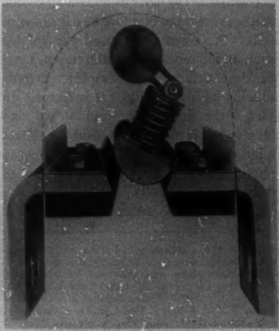


Fig. 5. Phantom view of new switch shown in open position

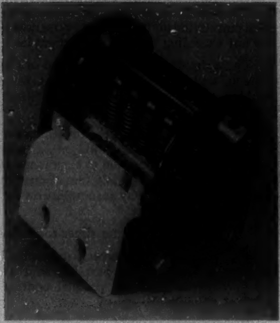


Fig. 6. New switch shown in closed position

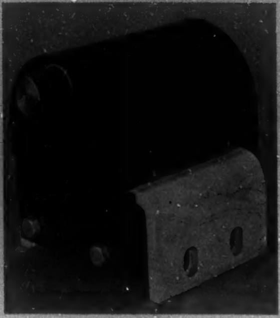


Fig. 7. New switch with plastic dust cover attached

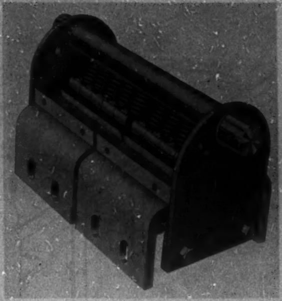


Fig. 8. Double unit of higher current rating than unit shown in Figs. 5, 6, and 7

Virginia, and Alabama. While they have proven satisfactory in service, the switches presented some manufacturing problems and were difficult to align during assembly. The contact surfaces on the rotating contact member were turned on a lathe and, hence, were readily held to close tolerances, but the amount of bend in the formed fingers tended to vary with the temper of the drawn copper. There were two contacts between the rotating beavertail and each of the contact fingers. It was found to be exceedingly difficult to keep the contact finger aligned so that both contacts on each finger were making after the terminals had been bolted to a bus bar.

Because of the experience with these switches, it was felt that an improved switch could be designed, and a list of the desirable characteristics of this switch was first set down on paper. The specifications were as follows:

1. The current path through the switch should be as short as possible to minimize the voltage drop across the switch and the material cost of the copper.

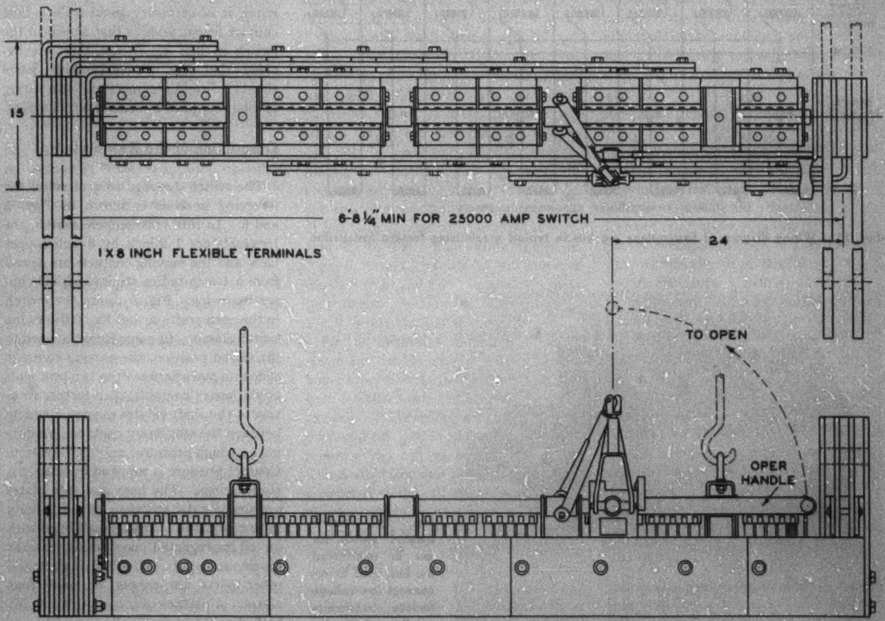


Fig. 9. 25,000-ampere portable switch assembly designed for use with diaphragm-type chlorine cell

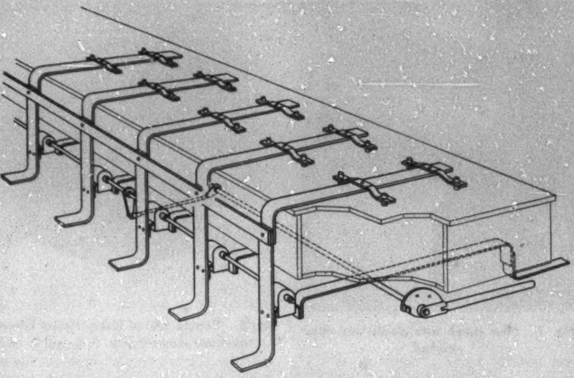


Fig. 10. The new switch applied to a typical mercury-type chlorine cell

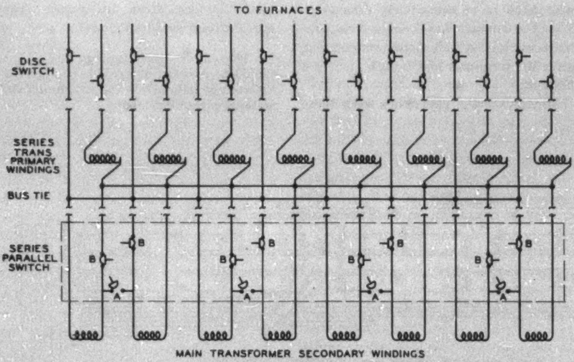

TO FURNACES
DISC SWITCH
SERIES TRANS PRIMARY WINDINGS
BUS TIE
SERIES PARALLEL SWITCH
B
A
MAIN TRANSFORMER SECONDARY WINDINGS

Fig. 11. Wiring diagram of low-voltage a-c bus in typical graphitizing furnace installation

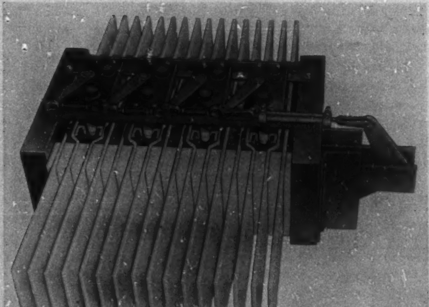


Fig. 12. Series-parallel switch for use in interleaved a-c bus, used to reconnect low-voltage furnace transformer windings in series and parallel groups

2. The switch should be able to carry its rated current indefinitely with a temperature rise not exceeding 30 degrees centigrade.

3. There should be ample deflection of the contacts under load so that contact wear would not appreciably reduce the contact pressure.

4. Contacts should be of the high-pressure limited-area type, and should make and break with a rubbing motion to insure metal-to-metal contact through any film of dust or foreign material that may accumulate on the contacts.

5. The contacts should be silver-to-silver.

6. The switch should have a number of independent contacts in parallel, so that there would be as many parallel contact points as possible.

7. The new switch should retain the feature of operation by 60-degree rotation of an operating shaft.

8. The switch parts should be adaptable to mass production methods to keep the cost to a minimum.

A study of specifications 1 and 2 indicates that the deflection of a spring member necessary to supply the contact pressure should be separated from the current-carrying parts. Because of the high currents involved, the current-carrying parts in a switch rated 3,000 amperes and above are necessarily quite heavy, and if there is to be a deflection of say 1/16-inch minimum, it is necessary to provide a long contact finger, such as that shown in the switch in Fig. 3. It was therefore decided that the current-carrying parts should be rigid members and the spring pressure supplied by some external spring.

Description of the New Switch

The switch designed as a result of the foregoing analysis is shown in Figs. 5 and 6. In this 3,000-ampere switch, the terminals are 1/2-inch by 6-inch copper bars, and the moving contacts are sawed from a 1-inch-radius semicylindrical copper extrusion. Fig. 5 shows the switch in the open position, and Fig. 6 shows the switch closed. In going from the open to the closed position, the moving contacts slide into place between the two terminals or stationary contacts, then further rotation of the shaft rotates moving contacts between the stationary contacts, producing a high-pressure rubbing motion. Contact pressure is supplied by stainless-steel springs. The base and side plates are steel for d-c switches, and aluminum for a-c switches. Bearings are lined with an oil-impregnated, nonferrous sintered bearing material. The shaft is steel, and other parts are copper or nonferrous alloys. A plastic cover can be readily attached as shown in Fig. 7.

When the switch is used for short-cir-

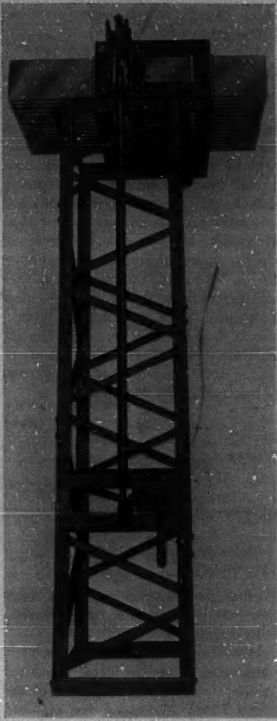


Fig. 13. Switch in Fig. 12 shown mounted on supporting steel rack, with operating mechanism installed. Bottom view of switch shows gear connections for simultaneous operation of contacts

cuiting electrolytic cells, it must, on opening, interrupt the full cell-line current at the cell voltage. While this voltage is not sufficient to maintain any arc, a sizable spark occurs when the contacts part, because of the concentration of current in a very small contact area. The contact surfaces fuse at this point, leaving a rough spot or crater which presents a poor contact surface when the switch is reclosed. Early designs with an arcing contact of a sintered refractory material such as silver-tungsten were discarded because the voltage drop across this contact was sufficient to cause burning at other contacts. In the final design, as seen in Fig. 7, all contacts are identical, but there is a spacer at the center contact in each group of five, so that this contact opens last and takes all the arcing. The switch is arranged so that the contact can be replaced readily when the surface becomes too badly pitted. Tests indicate that such replacement will be necessary about every 500 operations.

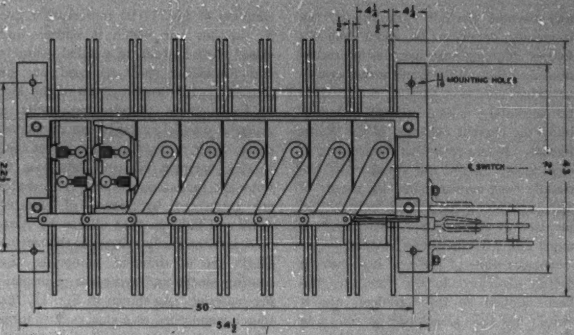


Fig. 14. Disconnecting switch for use in interleaved a-c bus, generally similar in design to the switch shown in Figs. 12 and 13

The voltage drop across the switch when carrying rated current is less than 20 millivolts. Maximum operating torque on the operating shaft is approximately 250 inch-pounds per 3,000-ampere switch.

This novel fundamental principle of design has made possible the construction of various special-purpose switches, and has found application as a cell short-circuiting switch, as a cell-disconnecting switch, as an a-c disconnecting switch, as a polarity-reversing switch for electroplating applications, and for switching connections between d-c generators and the anodes in electrolytic tin-plate lines. It has been applied satisfactorily at voltages up to 250. The switch shown in Figs. 5, 6, and 7 is rated 3,000 amperes direct current and 2,000 amperes alternating current. A double unit, as shown in Fig. 8, is rated at double these currents.

Fig. 9 illustrates a switch built of single and double units to form a 25,000-ampere portable switch intended for use with a well-known diaphragm-type chlorine cell. This switch is supported from an overhead rail, and is pushed into position opposite the cell to be short-circuited. The flexible terminal connectors are then connected to the cell short-circuiting terminals and the switch is closed. When repairs to the cell are finished, the switch is opened, and the terminals are disconnected and rotated to the vertical position.

Fig. 10 shows the switch applied to a typical mercury-type chlorine cell. The bus running from one chlorine cell to the next is usually divided into a number of parallel bars spaced about 3 to 5 feet apart, depending on the configuration of the anodes. One switch is mounted at each of these bars, as shown, in such a manner that one terminal is connected to the anodes and the other terminal to the cell bottom, which serves as the cathode. Where the total cell current is over 18,000 amperes, the operating handle is usually mounted on an auxiliary shaft with a toggle connection to keep the operating effort within reasonable limits.

Electric Furnace Switches

In graphitizing plants, the resistance of the furnace is relatively high at the beginning of the run, and a relatively high voltage is required. Furnace transformers are usually supplied with multiple low-voltage windings, as shown in Fig. 11. As the charge in the furnace heats up, its resistance decreases, and, to hold the current approximately constant at the maximum value for which the transformers and bus are designed, the voltage impressed across the furnace is decreased. This is usually accomplished by means of a load tap changer on the high-voltage side of the furnace transformer. However, in most cases, the range is too great to be covered by this means alone, and the low-voltage transformer windings are connected in series-parallel combinations to double the range of the primary tap changer. During the initial part of the run, the secondary windings are connected in parallel groups with two windings in series in each group. At this time, the switch contacts, Fig. 11(A), are closed, and the contacts, Fig. 11(B), are open. After the charge has heated up and become partially graphitized, the resistance decreases to about half its former value, and the switches are reconnected so that all the transformer secondary windings are in parallel, thus cutting the secondary voltage in half and doubling the current capacity of the transformer. In the past, this change-over of connections in the low-voltage bus has usually been accomplished by means of disconnecting links, which were bolted in place in either one of two positions, much like a double-throw switch.

The principle of the new high-current switch described in the foregoing has been applied to this problem, and a switch has been developed as shown in Fig. 12. Half of the bars on which no switching is necessary are run directly through the switch and kept in their true interleaved position, with respect to the switched bars. The reactance of the switch is therefore kept to a minimum, and the space requirements and cost are also minimized. Each of the contacts of Fig. 11(B) represents a gap in the bus bar which is opened and closed with a group of parallel contacts and associated shaft as used on the separate switch units. Each of the Fig. 11(A) contacts likewise represents a group of parallel contacts and associated shaft; these contacts make a direct connection between adjacent bus bars to form the series connection. The shafts are connected by means of suitable gears and links so that all can be operated simultaneously. Because the various shafts are at different potentials, the bearing blocks and gears are of insulating material. As shown in Fig. 13, the switch is mounted on a steel rack to form a free-standing unit, and all of the switch contacts are operated at the same time from a single hand crank.

Fig. 14 shows a similar switch arranged as a disconnecting switch to open and close a gap in all of the busses simultaneously. This switch is applicable wherever it is desired to isolate a portion of the interleaved a-c bus run from the live secondary bus.

Conclusion

The switch presented in this paper represents the culmination of nearly a decade of development. The novel concept of a short bridging contact, controlled from a rotatable shaft, opening and closing a wedge-shaped gap in a bus has made possible an entirely new type of switch. This switch is simple in design, rugged in construction, and reliable in operation. Within its voltage and current range, it is probably the most economical switch in commercial production at present. It is versatile in its application, being adaptable to single- or double-throw d-c or a-c switches, and can be mounted in any position.

◆

No Discussion

Linear Compensation of Saturating Servomechanisms

J. R. BURNETT
ASSOCIATE MEMBER AIEE

P. E. KENDALL
NONMEMBER AIEE

OPTIMUM switching of second-order contactor servomechanisms has been treated in the literature. A method of approximating this performance without the use of relays or computers consists of substituting a particular type of network and a limiter for the computer and relays. A procedure is given for obtaining the characteristics of this network. It is illustrated by examples and experimental results. This type of linear compensation yields results that are better than those obtained by conventional linear compensation having uncontrolled saturation.

Optimum switching of a contactor servomechanism has been treated extensively by means of the phase plane in recent years.[1,2] The philosophy of this treatment has been an application of full torque with a minimum number of torque reversals so as to reduce the error of the system to zero in a minimum time. It has been shown[1,2] that a second-order servo can be operated in this optimum manner by a single torque reversal. The control signal is normally a nonlinear function of the error and error rate, and usually requires a computer to determine the exact torque reversal point. It has also been shown in these references that the equation to be solved by the computer is

$$\left(\frac{dE}{dt}\right)^2 \pm 2\frac{T_m}{J}E = C \qquad (1)$$

where

$\frac{dE}{dt}$ = error rate
E = error
T_m = maximum torque of the control motor and is assumed to be constant
J = inertia of motor and load
C = constant

Further, for this equation, $C=0$, determines the point at which torque reversal must occur.

A contactor servomechanism suffers quite often from the disadvantage of having a dead or inoperative zone for small errors. McDonald[1] has proposed that a linear zone be substituted for this dead zone, and such a system has been called a dual-mode servomechanism. Hopkin[3] has used equation 1, but has fitted a linear zone around the complete curve specified by the equation when $C=0$.

This paper offers another solution to the problem of attaining this optimum

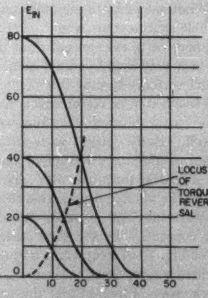


Fig. 1. Network input voltage

Paper 54-125, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 30, 1953.

J. R. Burnett is with Purdue University, Lafayette, Ind., and P. E. Kendall is with the Capehart-Farnsworth Corporation, Fort Wayne, Ind.

The work described in this paper was carried out as part of the nonlinear servomechanisms contract no. *AF 33(038)-21673* between Cook Research Laboratory and the United States Air Force.

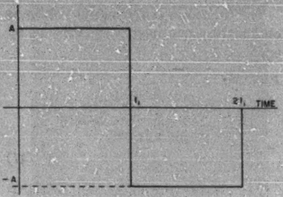


Fig. 2. Torque modulation

performance without the use of relays or computers. Modern network synthesis procedures are used to realize a linear network which operates upon the error signal and produces an output voltage that has zeros in time corresponding to the proper torque reversal points. This network then operates the control motor by means of a power amplifier. The network is followed by a limiter so that the voltage applied to the control motor will be a flat-topped wave. The problem has been limited to the case of a second-order servomechanism that has no output damping. It is believed that the method can be applied to more complex systems, but this will require determining their switching criteria.

The transfer function of a network specifies its performance, and the initial problem considered is that of obtaining the desired linear network transfer function by means of equation 2, since $E_{out}(s)$ and $E_{in}(s)$ are known. This transfer function is a linearized description of the ideal torque-switching occurring in a contactor servomechanism.

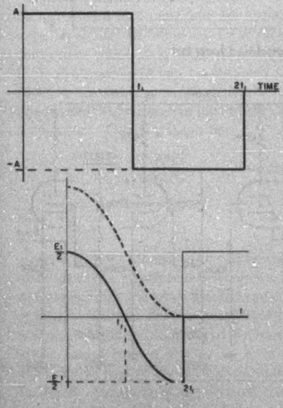


Fig. 3. Network output voltage

$$H(s)=\frac{E_{out}(s)}{E_{in}(s)} \tag{2}$$

$E_{out}(s)$ can assume several forms, and $E_{in}(s)$ is determined from a transient analysis of the optimumly switched contactor servomechanism. The inverse transform of $E_{in}(s)$ is pictured in Fig. 1. The synthesis of a network from an $H(s)$ is not a unique process; as a result, the final form of the network will depend very largely upon the designer.

Determination of *H*(*s*)

The determination of $H(s)$ is divided into the calculation of the Laplace transform of the input voltage and the calculation of the Laplace transform of the output voltage. The input voltage, since it is fixed by the system, is calculated first. The input voltage to the network consists of two sections of a parabola as shown in Fig. 1, and torque reversal occurs $t=t_1$. In choosing this input voltage it has been assumed that the system input is a step function of position, and that the system is operating as an optimumly switched contactor servomechanism. The known torque curve is shown in Fig. 2. The transform of the torque is given in equation 3

$$T(s)=\frac{T_m}{s}(1-2e^{-t_1 s}+e^{-2t_1 s}) \tag{3}$$

The torque is related to the output position by equation 4

$$L^{-1}[T(s)]=J\frac{d^2\theta_0}{dt^2} \tag{4}$$

The transform of equation 4 is

$$T(s)=J\left[s^2\theta_0(s)-s\theta_0(0+)-\left(\frac{d\theta_0}{dt}\right)\Big|_{0+}\right] \tag{5}$$

For a step input of position the following relations can be assumed: $\theta_0(0+)=0$ $(d\theta_0/dt)\big|_{0+}=0$, and $\theta_i(s)=E_1/s$.

In addition

$$\theta_0(s)=\theta_i(s)-E(s) \tag{6}$$

Combining equations 3, 5, and 6, and considering the initial conditions gives

$$E(s)=\frac{E_1}{s}-\frac{T_m}{Js^3}(1-2e^{-t_1 s}+e^{-2t_1 s}) \tag{7}$$

A study of the parabolic input yields[1]

$$\frac{T_m}{J}=\frac{E_1}{t_1^2} \tag{8}$$

With equation 8, equation 7 can be rewritten as

$$E(s)=\frac{E_1}{s}-\frac{E_1}{t_1^2 s^3}(1-2e^{-t_1 s}+e^{-2t_1 s}) \tag{9}$$

Equation 9 is the desired transform of the input voltage to the passive network.

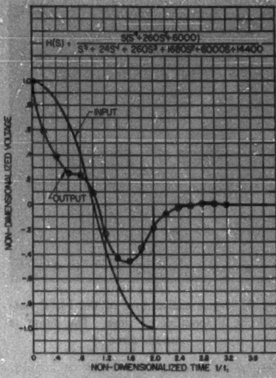


Fig. 4. Output and input of network

The transform of the output voltage is not as easily obtained since the requirements on the output are not all inclusive. It is necessary only that the output be zero at the torque reversal point $t=t_1$, at $t=2t_1$, and thereafter. Between zero time and $t=t_1$ a single polarity should exist, and between $t=t_1$ and $t=2t_1$ the opposite polarity should exist. The exact wave form during these intervals is unimportant, as a limiter will eliminate the irregularities and produce a square wave to drive the control motor into torque saturation.

Fig. 5. A 2-terminal pair network

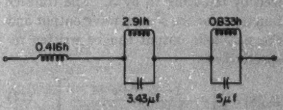


Fig. 6. Circuit element adjustment. Elements adjusted for $t_1=0.01$ second, impedance level IK

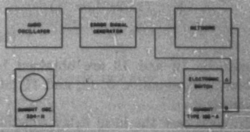


Fig. 7. Block diagram for open-loop testing

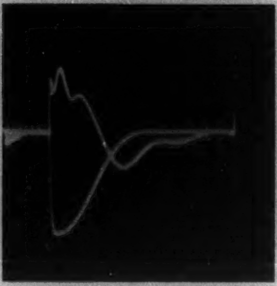


Fig. 8. Input and output of the $H_p(s)$ network as a function of time. The input is smooth curve

Vertical scale: arbitrary
Horizontal scale: 0.166 millisecond per large division

Two possible output wave forms, shown in Fig. 3, are considered. It should be realized, however, that infinitely many choices might be made. The transform of the square-wave output will be equivalent to equation 3 with T_m replaced by the amplitude of the output voltage A.

$$E_{out\,s}(s)=\frac{A}{s}(1-2e^{-t_1s}+e^{-2t_1s}) \tag{10}$$

The second form of output voltage considered is the form of the input displaced vertically by one-half the initial magnitude of the input. The transform for this function is

$$E_{out\,p}(s)=\frac{E_1}{2s}-\frac{E_1}{{t_1}^2s^3}(1-2e^{-t_1s}+e^{-2t_1s})+\frac{E_1}{2s}e^{-2t_1s} \tag{11}$$

The transform equations can be normalized by replacing st_1 by S. The transfer functions for the square-wave output and the parabolic output can be written respectively as

$$H_s(S)=\frac{A}{E_1}\left[\frac{1-2e^{-S}+e^{-2S}}{1-\frac{1}{S^2}(1-2e^{-S}+e^{-2S})}\right] \tag{12}$$

$$H_p(S)=\frac{\frac{1}{2}(1-e^{-2S})-\frac{1}{S^2}(1-2e^{-S}+e^{-2S})}{1-\frac{1}{S^2}(1-2e^{-S}+e^{-2S})} \tag{13}$$

Equations 12 and 13 are obtained by substituting equation 9, 10, and 11 into equation 2.

In each case the ideal delay function e^{-S} appears. This is inevitable because of the nature of both the input and output, and cannot be avoided. The ideal delay function is not physically realizable by passive lumped constant circuits; hence, an approximation is necessary since distributed constant circuits are not desired. The ideal delay function can be approximated by one of Pade's rational fraction approximations. The third approximant is

$$e^{-S}=\frac{1-\frac{1}{2}S+\frac{1}{10}S^2-\frac{1}{120}S^3}{1+\frac{1}{2}S+\frac{1}{10}S^2+\frac{1}{120}S^3} \tag{14}$$

A criterion of the value of the approximation will be the closeness of the actual wave form of the output voltage to the desired wave form. When this e^{-S} substitution is made in equations 12 and 13, the resulting transfer functions are, respectively

$$H_s(S)=\frac{A}{E_1}\times\left[\frac{4(S^4+120S^2+3{,}600)}{S^4+24S^3+200S^2+960S+3{,}300}\right] \tag{15}$$

$$H_p(S)=\frac{S(S^4+260S^2+6{,}000)}{S^5+24S^4+260S^3+1680S^2+6{,}000S+14{,}400} \tag{16}$$

The analytical output voltage of the parabolic transfer function is computed next as a test of the approximation.

The transform of the output voltage is obtained by multiplying equation 16 by equation 7.

$$E_{out}(S)=\frac{E_1t_1(S^4+260S^2+6{,}000)[S^2-(1-e^{-S})^2]}{S^2(S^5+24S^4+260S^3+1{,}680S^2+6{,}000S+14{,}400)} \tag{17}$$

The inverse transforms can be calculated by a partial fraction expansion of the total transform and by then writing the inverse transforms of the partial fractions. The problem is straightforward, but it involves the calculation of the roots of the 5th degree polynomial. The final calculations of the output are plotted in Fig. 4 where, also, the desired output of the network is shown for comparison. It is seen that the crossover occurs slightly late, as does the final zero point. Better results would be obtained by higher order approximations, but this would be at the expense of more elements in the network. Similar results are obtained by using the square-wave output transfer function.

Synthesis of the Network

The configuration of the network can have a variety of forms, depending upon the choices of the designer with regard to both the output voltage shape and the

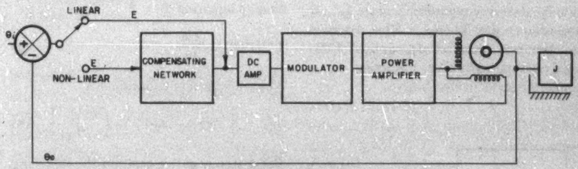


Fig. 9. Block diagram for closed loop test

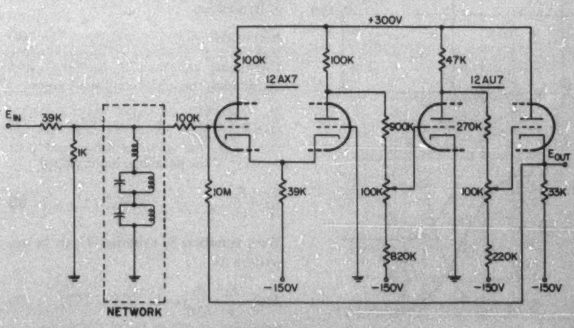


Fig. 10. Network and associated circuits

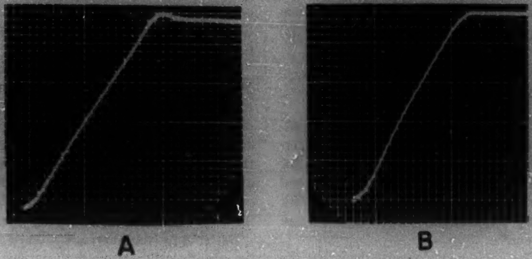


Fig. 11. Error as a function of time: 45-volt input step

Vertical scale: 18 volts per large division
Horizontal scale: 0.1 second per large division
A. For linearly compensated servomechanism
B. For servomechanism compensated with $H_p(s)$

synthesis procedure. The specific network considered for realization is the parabolic output one. Similar results are obtained for the square-wave output network. One configuration can be obtained if $H(s)$ is considered to be the transfer impedance function $Z_{12}(S)$ of a loaded network. This is illustrated in Fig. 5.

$$H_p(S)=Z_{12}(S)=\frac{E_2}{I_1}(S) \qquad (18)$$

The numerator of equation 16 or equation 18 is an odd function of S and the denominator is Hurwitz; consequently, the loaded transfer impedance can be written in terms of the unloaded-transfer impedance and the driving-point impedance, as seen from the output terminals

$$Z_{12}(S)=\frac{z_{12}(S)}{1+z_{22}(S)} \qquad (19)$$

The network at this point is assumed to be terminated in 1 ohm, but the impedance level can be adjusted to a more convenient value at a later time. The form of equation 19 can be obtained by dividing the numerator and denominator of equation 16 by the even part of the denominator. Then, by inspection

$$z_{12}(S)=\frac{S^5+260S^3+6{,}000S}{24(S^4+70S^2+600)} \qquad (20)$$

$$z_{22}(S)=\frac{S^5+260S^3+6{,}000S}{24(S^4+70S^2+600)} \qquad (21)$$

From equations 20 and 21 it can be seen that not only are both functions reactance functions; they are also a special type of a reactance function and can be realized by a shunt impedance, z_{22}, a shunt 1-ohm resistor, and a current source. A form of the reactive part of the network is shown in Fig. 6. Synthesis of the network from equations 20 and 21 will produce a network for normalized S so that the parameters of the final network must be adjusted by the normalizing factor t_1. Most servo systems have a value of t_1 that is less than 1 second, so the original element values obtained from the synthesis procedure will be reduced in value. The element values shown in Fig. 6 are based on a 1,000-ohm impedance level and a value for t_1 of 0.01 second. Since denormalization involves t_1 which is a function of E_1, the element values depend upon $\sqrt{E_1}$ also. For a restricted range of E_1 this variation will be slight.

Experimental Results

The network was tested experimentally by both open-loop and closed-loop methods. It is advisable first to consider the performance that is to be expected. First, with the choice of a value for t_1, the network parameters have been established for a particular magnitude of step-input. Any deviation from this value of input should not be expected to produce optimum performance, as the elements must vary with E_1. That a system will, in practice, operate satisfactorily over a range of inputs is demonstrated in the experimental testing; however, for large variations in E_1 it would be advisable to arrange for an adjustment of the parameters as a function of the input. Secondly, the system considered has zero damping. Therefore, for small perturbations around zero error some external damping should be added. In the experimental testing, a lead network was used for damping, both for the linear system and for the nonlinear system. It is necessary that any small signal corrective networks do not appreciably affect the zero crossings caused by the $H(s)$ network.

OPEN LOOP TESTING

It was necessary to have a function generator that could generate the parabolic wave form for open-loop testing. The first half-cycle of a cosine wave, displaced so as to be of one polarity, is a satisfactory approximate for the input voltage wave form and is more convenient to supply. The block diagram of the testing arrangement is shown in Fig. 7. A frequency of 1,000 cycles per second for the cosine wave was chosen to facilitate the use of the electronic switch. At a frequency of 1,000 cycles per second, the value of t_1 will be 0.25 millisecond, and the parameters of the network are smaller than those shown in Fig. 6. The results of the test are shown in Fig. 8. The input is represented by the negative wave form, and the output is represented by the initially positive wave form. The theoretical and experimental results both indicate a first zero crossing that is late. The trailing edge of the wave form returns to zero much later than it should.

CLOSED LOOP TESTS

One of the causes of the discrepancy in trailing-edge performance is the resistance of the inductances. However, the final proof of performance lies in the closed-loop performance, and the network was inserted in a closed loop as shown in Fig. 9.

Synthesis of the network was accomplished on the basis of a current source driving the network. The circuit of Fig. 10 was used for obtaining the current source and the effect of the network.

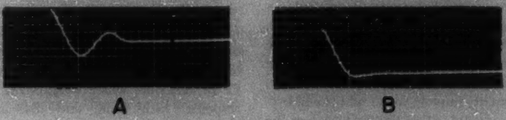


Fig. 12. Error as a function of time: 15-volt input step

Vertical scale: 30 volts per large division
Horizontal scale: 0.1 second per large division
A. For linearly compensated servomechanism
B. For servomechanism compensated with $H_p(s)$

To obtain optimum performance, the network parameters were experimentally adjusted. The experimental values did not differ greatly from the theoretical values. A certain adjustment is to be expected, since not only are the values of T_m and J used in the determination of t_1 not precise, but also the inductors are dissipative.

The performance of the nonlinear system using the network, followed by a limiter, was compared with the performance of a high-gain linear system. The results are shown in Figs. 11 and 12 as error-versus-time oscillograms. In Fig. 11 the network was adjusted for a step input of 45 volts, and with this same adjustment the input was changed to a step of 15 volts. The performance is better for the nonlinear system than for the linear system even with this change in input amplitude. For a more extensive change in amplitude, the elements of the $H(s)$ network would have to be changed as a function of the input. The linear system saturates at some level, but since attention is given the linear performance in compensating this system, the system having saturation controlled by the $H_p(s)$ network yields better performance.

Conclusions

The method presented here is an attempt to obtain optimum transient performance from a second-order contactor servo without the use of computers and relays. Relays are slow in operation for many applications, and computers are expensive.

The experimental results show that the method will produce responses which are an improvement over the same linear system even with the approximations used. The linear system saturates, but its criterion of design is small signal response. Little attention was given to the problem of small signal compensation in the controlled saturation system, but this can be handled by any of the usual procedures. In making the analytical calculations of the output wave form, the poles and zeros of the network were obtained so that Evans' root locus method is quite applicable, subject to the restriction of not changing the zero crossing times.

The design procedure is based on a prior knowledge of the ideal error and torque signals that exist for a step input of position. To apply the method to higher order systems it will be necessary to have the switching criterion for such a system available, and up to now the switching criterion[3] for systems of higher order than the third have not been completely solved.

Better approximations and attention to the theoretically required variations in network parameters should result in improved transient performance.

References

1. MULTIPLE MODE OPERATION OF SERVOMECHANISMS, D. McDonald. *Review of Scientific Instruments*, New York, N. Y., vol. 23, no. 1, Jan. 1952, pp. 22–30.

2. A PHASE-PLANE APPROACH TO THE COMPENSATION OF SATURATING SERVOMECHANISMS, Arthur M. Hopkin. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 631–39.

3. AN INVESTIGATION OF THE SWITCHING CRITERIA FOR HIGHER ORDER CONTACTOR SERVOMECHANISMS, I. Bogner. *Ph.D. Thesis*, University of Michigan, Ann Arbor, Mich., 1953.

No Discussion

The Transient Performance of Servomechanisms with Derivative and Integral Control

RICHARD C. LATHROP
ASSOCIATE MEMBER AIEE

DUNSTAN GRAHAM
MEMBER AIEE

THE utility of servomechanisms in their many applications depends on the speed, stability, and accuracy of their responses. These three characteristics, or the lack of them, may be observed in the responses of a servomechanism to test inputs. The error, which is the difference between the input and the output of the device, is used to actuate the control in a direction which tends to remove the error. In order to achieve the best performance under both transient and equilibrium conditions, it often has been found expedient to employ, as an additional correction signal, either the derivative or the time integral of the error, or both. The servomechanism is then said to have derivative and/or integral control. The present paper reports an investigation of the transient performance of low-order, linear servomechanisms with derivative and integral control.

In a previous paper[1] the authors developed a criterion of transient performance for servomechanisms based on the minimization of the integral of time-multiplied absolute-value of error (ITAE) $\int_0^\infty t|e|dt$. Factors affecting the generic qualities of speed, stability, and accuracy of the transient responses of servomechanisms are taken into account by the ITAE criterion, which weights the error response with time, and disapproves equally both positive and negative errors. Throughout the paper the term "optimum system" refers to the system of a given form which produces a minimum ITAE for a specified test input. Normalized transfer functions corresponding to the optimum system configurations of servomechanisms employing derivative and integral signals have been discovered and are tabulated for use in design.

The test inputs most commonly employed in mathematical or experimental transient analyses of servomechanisms are the impulse, step function, ramp function (velocity step), and the acceleration step. Each of these functions is, of course, the time integral of the preceding one, and a similar relation obtains between the responses which they produce from a linear servomechanism. For this reason, and since in linear systems the magnitude of the response is proportional to the magnitude of the input, attention is often concentrated on the error response to a unit displacement step input and on the steady-state errors in following unit steps of velocity and acceleration.

It is commonly desirable to have zero steady-state displacement error following

---

Paper 54-126, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing November 25, 1953.

RICHARD C. LATHROP and DUNSTAN GRAHAM are with the Wright Air Development Center, Wright Patterson Air Force Base, Ohio.

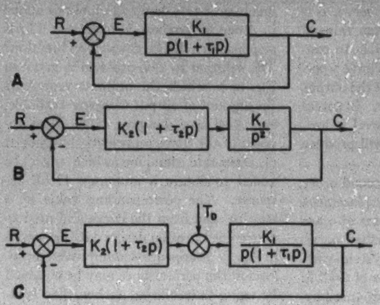


Fig. 1 (left). Block diagrams of second-order systems

A—Basic duplicator with output damping
B—Duplicator with error-rate damping
C—Duplicator or regulator with error-rate and output damping

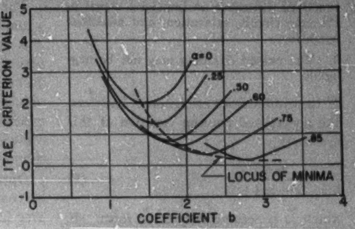


Fig. 2. The ITAE criterion applied to the displacement step-function responses of second-order systems with the transfer function

$$\frac{C(s)}{R(s)}=\frac{\alpha bs+1}{s^2+bs+1}$$

a displacement input, and to insure that the steady-state errors in following velocity and acceleration inputs are zero or as small as possible. In terms of the general, normalized transfer function[1] for servomechanisms

$$\frac{C(s)}{R(s)}=\frac{p_m s^m+p_{m-1}s^{m-1}+\ldots+p_2s^2+p_1s+p_0}{s^n+q_{n-1}s^{n-1}+\ldots+q_2s^2+q_1s+1} \tag{1}$$

the steady-state error may be shown to be

$$e_{ss}=\underset{t\to\infty}{e(t)}=(p_0-1)r+(p_1-q_1)\frac{dr}{dt}+(p_2-q_2)\frac{d^2r}{dt^2} \tag{2}$$

Zero steady-state displacement error is therefore achieved when $p_0=1$, and small velocity error is achieved by making the difference p_1-q_1 small. Integral control may be applied to accomplish these effects.

Nomenclature

$r(t)$=input to a servo system
$c(t)$=output, or response, of a servo system
$e(t)=r(t)-c(t)$=servo system error
$E(p)$=Laplace transform of error
$E(s)$=normalized Laplace transform of error
e_{ss}=steady-state error
t=real time
$\tau=\omega_0 t$=normalized time
p=Laplace complex variable
s=normalized Laplace variable
ω_0=natural angular frequency
τ_i=system time constant
α=a system constant
$a, b, c \ldots$=normalized transfer function coefficients
p_i, q_i=normalized transfer function coefficients
ζ_1=damping ratio due to output damping
ζ_2=damping ratio due to error-rate damping
ζ=total damping ratio
T_D=transform of load disturbance
K_i=a system gain parameter
K=total static loop gain

Second-Order Systems

A duplicator servomechanism is one in which the output motion approximately reproduces the shape of the input motion. One of the simplest mathematical models of a duplicator servomechanism is the one given by the block diagram of Fig. 1(A). Ideally, such a system may be described by a linear second-order differential equation with constant coefficients. An equivalent and more useful description of the system is obtained from the system transfer function

$$\frac{C(p)}{R(p)}=\frac{\omega_0^2}{p^2+b\omega_0 p+\omega_0^2} \tag{3}$$

The constants b and ω_0 are related to the system parameters of Fig. 1(A) by the expressions $b=1/(\omega_0\tau_1)$ and $\omega_0^2=K_1/\tau_1$.

The functional forms of Fig. 1(A) and equation 3 show a term proportional to the rate of change of the output. In practice, this so-called output damping may appear in one of the following three forms:

1. Direct damping of the output quantity, such as might be caused by viscous friction.
2. Indirect, or internal, damping of the output quantity, such as the effect of armature resistance in a d-c servomotor.
3. Feedback damping, such as might be obtained by adding a tachometer generator signal to the input of a servomotor.

The three forms are mathematically identical. An important practical difference, however, is that the first type absorbs power, while the others do not. Nevertheless, for the purposes of this paper, the three types will be considered equivalent, and will be classified collectively as output damping. The second-order servomechanism employing only output damping will be termed the basic system.

Using the change of variable $p=\omega_0 s$, we can reduce equation 3 to the form

$$\frac{C(s)}{R(s)}=\frac{1}{s^2+bs+1} \tag{4}$$

This normalization procedure is equivalent to a linear change of time scale, and is applicable to similar systems of higher order. The number of coefficients in the denominator of the transfer function is reduced to one less than the order of the system, and the coefficients of the first and last terms of the denominator become unity. The normalization technique is described in Appendix I of reference 1.

Another mathematical model of a simple servomechanism is the one illustrated in Fig. 1(B). The physical embodiment of this model might consist of an instrument servomechanism with negligible output damping, and with an error-rate, or derivative control circuit added to provide damping. In practice, an error-rate-damped servo may differ from the mathematical model in the following respects:

1. It may be difficult to achieve reasonably pure differentiation.
2. A rapidly changing input, such as a step

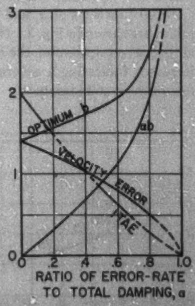


Fig. 3. Performance of the optimum second-order displacement duplicators

function, will cause an impulsive signal in the rate circuit, saturation, and nonlinear operation.

3. The output damping may not be negligible.

If these difficulties are ignored, however, the normalized transfer function of this system is

$$\frac{C(s)}{R(s)}=\frac{bs+1}{s^2+bs+1} \tag{5}$$

which is a zero-velocity-error form. If equation 5 is written as

$$\frac{C(s)}{R(s)}=\frac{1}{s^2+bs+1}+\frac{bs}{s^2+bs+1} \tag{6}$$

the step-function response of the error-rate-damped system may be seen to be equal to the step function response plus b times the weighting function[2] of the system with equivalent output damping. It is clear, therefore, that the step-function responses of two systems with equal damping may be considerably different.

The output-damped and error-rate-damped second-order servomechanisms are each special cases of a more general system employing both types of damping. The block diagram of such a system is shown in Fig. 1(C), and the normalized transfer function is

$$\frac{C(s)}{R(s)}=\frac{\alpha bs+1}{s^2+bs+1} \tag{7}$$

The quantity α is the ratio of error-rate damping to total system damping, and b represents the total system damping. If the system is to be used to follow sudden changes of input displacement, then the unit input displacement step function is a suitable test signal with which to assay system performance. In this case the ITAE as a function of b for various given values of α is given by the curves of Fig. 2. The minimum values of the ITAE, the steady-state velocity error of the servo system, and the optimum values of b and αb are all functions of α, and these relationships are shown in Fig. 3. It may be noted that, as α is increased, the ITAE value decreases, indicating a general improvement in the system response.

For the zero-velocity-error servo, α is equal to unity, and the optimum value of b is infinite. If it were possible to construct such a system, it would have a perfect response, in the sense that the output would follow the input exactly. In actual practice, if the input noise level is low, values of b greater than 3.0 will produce very good responses.

Optimum responses of the second-order servomechanisms to input displacement step functions for various values of α are presented in Fig. 4. The dashed curve shows the response of the system with $\alpha=1$ and a compromise value of $b=3.2$. The responses of Fig. 4 were produced by systems for which the ITAE was minimized for displacement inputs, and one might reasonably inquire into the behavior of these servos with velocity step-function inputs. Fig. 5 shows the error responses of these optimum-displacement systems with unit velocity inputs.

For the zero-velocity-error system ($\alpha=1$), the error signal following a unit velocity input can be used in the ITAE criterion to obtain the optimum system adjustment for velocity inputs. If this is done, the value $b=1.7$ is found to be optimum and the corresponding error response has the form of curve A, Fig. 7.

The discussion so far has assumed that the system parameters were completely at the designer's disposal. An interesting and practical problem in servo design might be stated as follows: given a second-order servomechanism with a certain amount of output damping, how much error-rate damping should be added in order to produce a minimum ITAE response to a displacement step function? The solution to this problem is given in Fig. 6, in which the ordinate represents the amount of output damping available, expressed as a damping ratio. The abscissa of the curve represents the amount of error-rate damping which must be added to obtain a minimum ITAE response. The corresponding value of α may be read from the curve and used to determine the other design data presented in Fig. 3. It appears that improved servomechanism performance may be obtained by using destabilizing tachometric feedback, and then making up the deficiency in stability with error-rate damping. This would result in a smaller amount of effective output damping, a smaller velocity error, and a general improvement in response to a displacement step function.

Servo systems may be designed to operate as regulators, instead of or in addition to operating as duplicators. Generally speaking, a good regulator is one which presents a high impedance to external disturbances, and which recovers from the effects of such disturbances quickly and accurately. The block diagram of Fig. 1(C) can serve as a model of the second-order regulator if the input $r(t)$ is assumed constant and the disturbance T_D is applied as indicated. As in the case of the duplicator, the optimum adjustment of the regulator depends upon the form of the disturbing quantity. Unlike the duplicator response, however, the regulator

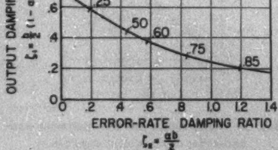


Fig. 4 (left). The error-rate damping ratio to add for optimum response

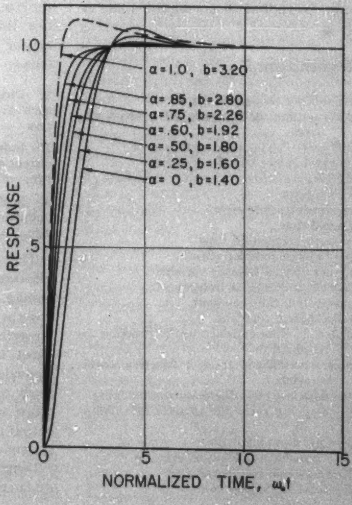


Fig. 5 (right). Displacement step-function responses of the optimum second-order duplicators

Table I. Second-Order Systems

Case No.	Description of the System	Fig. No.	Type of Damping	Assumed Input Function	Remarks	Normalized Transfer Function	Normalized Transform of Error	Literal Relations Between Coefficients	Optimum Coeff. as Selected by ITAE	Fig. No. of Optimum Response
1	Basic Second-order Servomechanism	1a	Output	Unit Displ. Step $R(s) = 1/s$	If a velocity step is applied, this system displays a steady-state error.	$\frac{C(s)}{R(s)} = \frac{1}{s^2 + bs + 1}$	$E(s) = \frac{s + b}{s^2 + bs + 1}$	$b = \frac{1}{\omega_o \tau_1}$ $\omega_o^2 = \frac{K_1}{\tau_1}$	$b = 1.4$	Fig. 5 ($\alpha = 0$)
2	Zero-velocity-Error Servo	1b	Error-Rate	Unit Displ. Step		$\frac{C(s)}{R(s)} = \frac{bs + 1}{s^2 + bs + 1}$	$E(s) = \frac{s}{s^2 + bs + 1}$	$b = \omega_o \tau_2$ $\omega_o^2 = K = K_1 K_2$	Ideally, $b = \infty$. Practical $b > 3.0$	Fig. 5 ($\alpha = 1$, $b = 3.2$)
3	Zero-velocity-error Servo	1b	Error-Rate	Unit Velocity Step $R(s) = 1/s^2$	Same system as Case 2, but adjusted for optimum response to velocity inputs.	$\frac{C(s)}{R(s)} = \frac{bs + 1}{s^2 + bs + 1}$	$E(s) = \frac{1}{s^2 + bs + 1}$	$b = \omega_o \tau_2$ $\omega_o^2 = K = K_1 K_2$	$b = 1.7$	Fig. 7a
4	Reduced-velocity-error Servo	1c	Output and Error-Rate	Unit Displ. Step $R(s) = 1/s$		$\frac{C(s)}{R(s)} = \frac{\alpha bs + 1}{s^2 + bs + 1}$	$E(s) = \frac{s + b(1-\alpha)}{s^2 + bs + 1}$	$b = \frac{1 + K\tau_2}{(K\tau_1)^{1/2}}$ $K = K_1 K_2$ $\omega_o^2 = K/\tau_1$ $\alpha = \frac{K\tau_2}{1 + K\tau_2}$ $\alpha b = \omega_o \tau_2$	Depends upon α. See Fig. 3	Fig. 5
5	Regulator	1c	Any Type or Types	Unit Impulse of Load $T_D(s) = 1$		$\frac{E(s)}{T_D(s)} = \frac{1}{K(s^2 + bs + 1)}$	$E(s) = \frac{1}{K(s^2 + bs + 1)}$	$b = \frac{1 + K\tau_2}{(K\tau_1)^{1/2}}$ $K = K_1 K_2$ $\omega_o^2 = K/\tau_1$	$b = 1.7$ K = large as possible	Fig. 7a
6	Regulator	1c	Any Type or Types	Unit Initial Condition of Error	Initial error created by forcibly displacing the output and then suddenly releasing.		$E(s) = \frac{s + b}{s^2 + bs + 1}$	$b = \frac{1 + K\tau_2}{(K\tau_1)^{1/2}}$ $K = K_1 K_2$ $\omega_o^2 = K/\tau_1$	$b = 1.4$	Fig. 7b

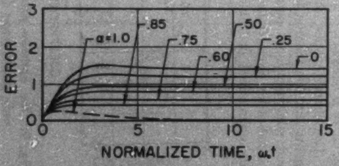
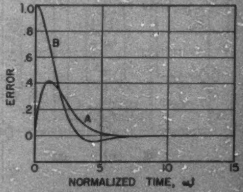


Fig. 6 (left). Error responses to unit velocity inputs of second-order duplicators optimized for displacement inputs

Fig. 7 (right). Error responses of the optimum second-order regulators

A—The unit output disturbance impulse response, b=1.7
B—The unit initial condition response, b=1.4



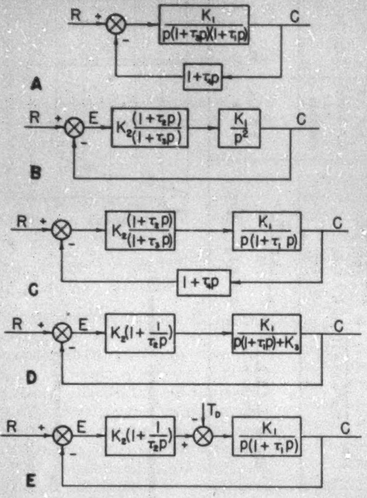

R
+
−
$\frac{K_1}{p(1+\tau_3 p)(1+\tau_1 p)}$
C
$1+\tau_4 p$
A
R
+
−
E
$K_2\frac{(1+\tau_2 p)}{(1+\tau_3 p)}$
$\frac{K_1}{p^2}$
C
B
R
+
−
$K_2\frac{(1+\tau_2 p)}{(1+\tau_3 p)}$
$\frac{K_1}{p(1+\tau_1 p)}$
C
$1+\tau_4 p$
C
R
+
−
E
$K_2(1+\frac{1}{\tau_2 p})$
$\frac{K_1}{p(1+\tau_1 p)+K_3}$
C
D
R
+
−
E
$K_2(1+\frac{1}{\tau_2 p})$
−
T_D
+
$\frac{K_1}{p(1+\tau_1 p)}$
C
E

Fig. 8 (left). Block diagrams of third-order systems

A—Basic duplicator with output damping
B—Duplicator with error-rate damping
C—Duplicator or regulator with error-rate and output damping
D—Duplicator with spring load and integral control
E—Duplicator or regulator with output damping and integral control

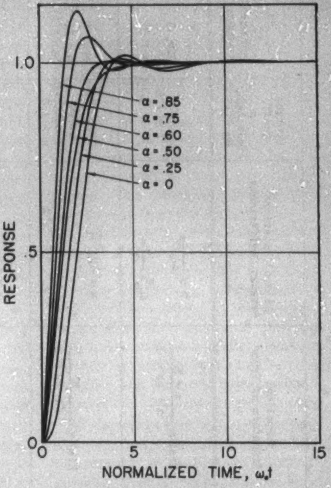


Fig. 10 (right). Displacement step-function responses of the optimum third-order displacement duplicators

response is independent of the type of damping which is used.

Two important disturbance test functions for second-order regulators for which the ITAE criterion converges are: 1. the unit impulse, and 2. the initial condition of error, created by forcibly displacing the output and then suddenly releasing it at $t=0$. The error response of the regulator to an impulsive disturbance is dynamically identical to that of a zero-velocity-error servomechanism with a velocity input, and the response corresponding to optimum adjustment of a system is given by curve *A* of Fig. 7. The error response of the regulator to a unit initial condition of error is dynamically identical to that of an output-damped servomechanism with a unit displacement input, and the optimum response is shown by curve *B* of Fig. 7.

The characteristics, optimum adjustments, and the relation between system and nondimensional parameters of the second-order duplicators and regulators are summarized in Table I.

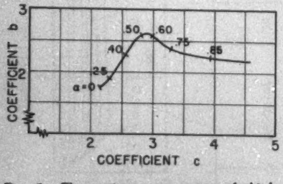


Fig. 9. The optimum parameters of third-order displacement duplicators

Third-Order Systems

The prototype model of a third-order duplicator servomechanism is given by the block diagram of Fig. 8(A). It may be noted that this model differs from the basic second-order model only in the presence of an additional lag in the forward loop. The normalized transfer function of the third-order prototype is

$$\frac{C(s)}{R(s)}=\frac{1}{s^3+bs^2+cs+1} \tag{8}$$

The model employs output damping, and exhibits a steady-state error proportional to the constant *c*, following a step velocity input. If the output damping is negligible, and a resistance-capacitance lead network is inserted in the forward loop for stabilization purposes, a system of the form of Fig. 8(B) is the result. (The denominator time constant τ_3 of the lead network is assumed to be large enough so that it is not negligible.) Such a system is characterized by the normalized transfer function

$$\frac{C(s)}{R(s)}=\frac{cs+1}{s^3+bs^2+cs+1} \tag{9}$$

This servomechanism exhibits no steady-state error following an input velocity step function, and, in this, it is analogous to the second-order error-rate-damped system. The two duplicators just described are again special cases of a more general servomechanism employing both a lead network and output damping, shown in Fig. 8(C). Such a system has the normalized transfer function

$$\frac{C(s)}{R(s)}=\frac{\alpha cs+1}{s^3+bs^2+cs+1} \tag{10}$$

There are also two specialized duplicators which employ integral control, and which have normalized transfer functions which reduce to the form of equation 10. One of these is a servomechanism with inherent output damping. Integral control is provided to eliminate the steady-state velocity errors caused by the output damping. The block diagram of such a system is illustrated in Fig. 8(E). The other third-order integral control duplicator servo, represented by Fig. 8(D), utilizes the integral control to eliminate steady-state displacement errors caused by a spring load on the output.

The optimum values of the coefficients *b* and *c* in equation 10 depend upon the value of α and the type of input. For systems in which the input is a displacement step function, the optimum values of *b*

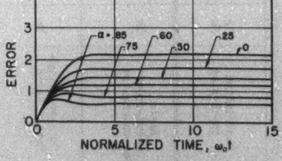


Fig. 11. Error responses to unit velocity inputs of third-order duplicators optimized for displacement inputs

Table II. Third-Order Systems

Case No.	Description of the System	Fig. No.	Type of Damping	Assumed Input Function	Remarks	Normalised Transfer Function	Normalised Transform of Error	Literal Relations Between Coefficients See Fig. 8	Optimum Coeff. as Selected by ITAE	Fig. No. of Optimum Response
1	Basic Third-order Servomechanism	8a	Output, velocity Feedback	Unit Displ. Step $R(s) = 1/s$	Formed from basic second-order system by an additional lag in the forward loop.	$\frac{C(s)}{R(s)} = \frac{1}{s^3 + bs^2 + cs + 1}$	$E(s) = \frac{s^2 + bs + c}{s^3 + bs^2 + cs + 1}$	$b = \frac{\omega_o^2(\tau_1 + \tau_2)}{K_1}$ $c = \frac{1}{\omega_o^2 \tau_1 \tau_3} + \omega_o \tau_4$ $\omega_o^3 = K_1/\tau_1 \tau_3$	b = 1.75 c = 2.15	Fig. 10 ($\alpha = 0$)
2	Zero-velocity-error system	8b	Error-rate	Unit Displ. Step $R(s) = 1/s$	Lead network in the forward loop provides total system damping.	$\frac{C(s)}{R(s)} = \frac{cs + 1}{s^3 + bs^2 + cs + 1}$	$E(s) = \frac{s^2 + bs}{s^3 + bs^2 + cs + 1}$	$b = \frac{\omega_o^2}{K} = \frac{1}{\omega_o \tau_3}$ $c = \omega_o \tau_2 \quad K = K_1 K_2$ $\omega_o^3 = K/\tau_3$	b = 1.75 c = ∞ (Practical) c ≥ 3.0	Fig. 24, Ref. 1 ($\alpha = 1$, $b = 1.75$, $c = 3.25$)
3	Zero-velocity-error system	8b	Error-rate	Unit Velocity Step $R(s) = 1/s^2$	Same system as Case 2, but adjusted for optimum response to velocity inputs.	$\frac{C(s)}{R(s)} = \frac{cs + 1}{s^3 + bs^2 + cs + 1}$	$E(s) = \frac{s + b}{s^3 + bs^2 + cs + 1}$	Same as Case 2	b = 1.30 c = 2.94	Fig. 12 Curve A
4	Reduced-velocity-error System	8c	Output plus Error-rate and velocity feedback	Unit Displ. Step $R(s) = 1/s$	Lead network inserted to improve system stability. Velocity feedback not required for $\alpha \geq .44$.	$\frac{C(s)}{R(s)} = \frac{\alpha cs + 1}{s^3 + bs^2 + cs + 1}$ $0 \leq \alpha \leq 1$	$E(s) = \frac{s^2 + bs + c(1 - \alpha)}{s^3 + bs^2 + cs + 1}$	$b = \frac{\tau_1 + \tau_3 + K\tau_2\tau_4}{\omega_o \tau_1 \tau_3}$ $\alpha = \frac{K\tau_2}{1 + K(\tau_2 + \tau_4)}, c = \omega_o \tau_2/\alpha$ $\omega_o^3 = K/\tau_1 \tau_3 \quad K = K_1 K_2$	Depends upon α. See Fig. 9	Fig. 10
5	Reduced-velocity-error servo with spring load and integral control	8d	Output	Unit Displ. Step $R(s) = 1/s$	Integral control removes steady-state error in displacement due to spring load.	$\frac{C(s)}{R(s)} = \frac{\alpha cs + 1}{s^3 + bs^2 + cs + 1}$	$E(s) = \frac{s^2 + bs + c(1 - \alpha)}{s^3 + bs^2 + cs + 1}$	$b = 1/\omega_o \tau_1 \quad \omega_o^3 = K/\tau_1 \tau_2$ $c = \frac{\omega_o \tau_2}{\alpha} \quad \alpha = \frac{K}{K + K_3}$	Depends upon α. See Fig. 9	Fig. 10
6	Zero-velocity-error system	8e	Output	Unit Velocity Step $R(s) = 1/s^2$	Integral control removes steady-state velocity error due to output damping.	$\frac{C(s)}{R(s)} = \frac{cs + 1}{s^3 + bs^2 + cs + 1}$	$E(s) = \frac{s + b}{s^3 + bs^2 + cs + 1}$	$b = 1/\omega_o \tau_1 \quad \omega_o^3 = K/\tau_1 \tau_3$ $c = \omega_o \tau_2$ $K = K_1 K_2$	b = 1.30 c = 2.94	Fig. 12 Curve A
7	Regulator	8e	Any type or types	Unit Initial Condition of error	Initial error created by forcibly displacing the output and then suddenly releasing.		$E(s) = \frac{s^2 + bs + c}{s^3 + bs^2 + cs + 1}$	$b = \frac{\tau_1 + \tau_3}{\omega_o \tau_1 \tau_3} \quad K = K_1 K_2$ $c = \frac{1 + K\tau_2}{\omega_o^2 \tau_1 \tau_3} \quad \omega_o^3 = K/\tau_1 \tau_3$	b = 1.75 c = 2.15	Fig. 12 Curve B
8	Regulator	8e	Any type or types	Unit Step of Load $T_D(s) = 1/s$	Integral control is used to eliminate steady-state displ. error due to load step.	$\frac{E(s)}{T_D(s)} = \frac{cs}{K_2(s^3 + bs^2 + cs + 1)}$	$E(s) = \frac{c}{K_2(s^3 + bs^2 + cs + 1)}$	Same as Case 6	b = 1.97 c = 2.41 K_2 = large as possible	Fig. 12 Curve C

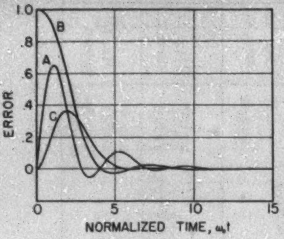


Fig. 12. Error responses of the optimum third-order systems with integral control

A—Optimum error response to a velocity input

B—Optimum initial condition response of the regulator

C—Optimum response of the regulator to a unit step function of output disturbance

and c, as selected by the ITAE criterion, are given in Fig. 9. The rather sharp reversal in the tendency of the criterion to pick higher and higher b values corresponds to a rejection of an overdamped response with a persistent error, and a switch in the responses which the criterion prefers to ones with a tendency toward a large first overshoot. This is illustrated in Fig. 10, which presents the unit step-function responses corresponding to various points along the curve of Fig. 9.

The systems defined by the coefficients given in Fig. 9 are designed to yield optimum responses to displacement step functions. If these same systems are subjected to velocity-step-function inputs, they will exhibit the error responses of Fig. 11, which are the integrals of the error responses of Fig. 9. A relatively large value of α, corresponding to a small difference c-αc, is seen to result in a small steady-state velocity error.

It is necessary to point out that the time constants and gain parameters of the various duplicators, all described by equation 10, enter into the coefficients b, c, and α in quite a different way in each case. Nevertheless, the selections of optimum parameters illustrated in Fig. 9 are valid for the synthesis of all of the third-order displacement duplicators described in the foregoing. The mathematical relationships between the time constants and gain parameters of the physical systems and the normalized representation of equation 10 are given in Table II.

For those systems which reduce to the zero-velocity-error form of equation 9 (systems in which $\alpha=1$, the system of Fig. 8(C), for example) it is possible to apply the ITAE criterion to the selection of optimum coefficients for velocity-step-function inputs. If this is done, the optimum error response takes the form of curve *A* of Fig. 12. This response is that of a system in which $b=1.30$ and $c=2.94$, values which are appreciably different from any along the curve of Fig. 9.

The general remarks about second-order regulators concerning the impedance presented to load disturbances apply as well to the third-order systems. One important regulator response is the one which occurs when the output of the servo of Fig. 8(C) is forcibly displaced and then suddenly released. This initial-condition response is shown, for the optimum case, by curve *B* of Fig. 12. This error response is dynamically identical to the step-function response of the duplicator described by equation 10 when $\alpha=0$, and the same values for the optimum coefficients, $b=1.75$, $c=2.15$, apply.

Another regulator configuration is presented in Fig. 8(E). If a load disturbance step function is applied to this regulator, the integral control reduces the steady-state error to zero. The error response for optimum system adjustment is given by curve *C* of Fig. 12.

Table II contains a summary of the descriptions, synthesis data, and optimum adjustments of the third-order duplicators and regulators, together with appropriate references to the block diagrams and the graphical representations of the responses.

The Response of Zero-Velocity-Error Systems

It has been noted that the error of a servomechanism in response to a velocity step input is the time integral of its error response to a displacement step input. This is in fact illustrated by the correspondence between Figs. 5 and 6 and between Figs. 10 and 11. Porter[3] has emphasized and illustrated the fact that the response to a step function of input displacement of a zero-velocity-error servomechanism is characterized by at least one overshoot. This is in order that the integral of error may approach zero. He has suggested that these observations may be useful in determining by experimental procedures the nature of the dynamic lags or forced errors[4] in a system.

It is also pertinent to say that the requirement on a zero-velocity-error servomechanism that the integral of its error response to a step function of input displacement approach zero precludes this response from having characteristics in accord with intuitive concepts of an optimum (small peak overshoot and no persistent errors, for example). This last fact is illustrated in Fig. 12 which presents three selected responses of third-order zero-velocity-error systems. The range of choices is from responses with a large transient peak enduring for a short period of time to ones with a small peak, but with a persistent mode. All three responses have equal amounts of integrated positive and negative error. A similar but aggravated situation with respect to the selection of an optimum obtains in connection with zero-acceleration-error systems.

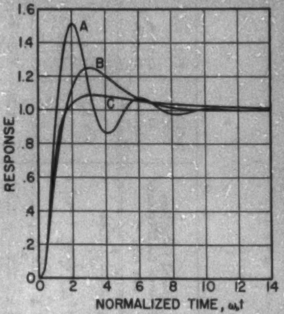


Fig. 13. The displacement step-function response of zero-velocity-error servo systems

A—$\mathcal{L}^{-1}\left[\frac{1}{s}\frac{2.94s+1}{s^3+1.30s^2+2.94s+1}\right]$ is the response of the optimum system

B—$\mathcal{L}^{-1}\left[\frac{1}{s}\frac{3s+1}{s^3+3s^2+3s+1}\right]$ is the response of the binomial system

C—$\mathcal{L}^{-1}\left[\frac{1}{s}\frac{6.3s+1}{s^3+5.1s^2+6.3s+1}\right]$ is the response of Whiteley's standard form system

When the ITAE criterion is applied to the selection of an optimum displacement step-function response of a zero-velocity-error system, it tends to select a response with a large overshoot rather than a persistent error. If the criterion is applied to the selection of an error response to a step velocity input, the corresponding displacement step function response exhibits a still larger overshoot. Such a response is shown in Fig. 13 as curve *A*. This is the displacement step-function response of the system whose error response to a velocity step is shown as curve *A* in Fig. 12.

Conclusion

In a previous paper, the authors developed the ITAE criterion and applied it to the selection of optimum normalized transfer functions (standard forms) for duplicators of the zero-displacement-error, zero-velocity-error, and zero-acceleration-error types. Only the transient responses to input displacement step functions were considered. In the present paper, the work has been extended to the types of second- and third-order duplicator systems which do not fall in one of the limiting case categories, to considerations of velocity inputs to duplicators, and to the response of regulators. The codified and tabulated results demonstrate that the optimum synthesis of servo systems is strongly dependent on the combination of types of damping which are employed, and on the input or disturbance on the output which is considered most typical of the intended application. Actual designs will probably represent compromises between various optima, but the tabulation of the most desirable parameter values for given conditions will prove to be a useful guide to the best compromise.

References

1. The Synthesis of "Optimum" Transient Responses: Criteria and Standard Forms, Dunstan Graham, R. C. Lathrop. *AIEE Transactions*, vol. 72, pt. II, Nov. 1953, pp. 273–88.

2. Theory of Servomechanisms (book), H. M. James, N. B. Nichols, R. S. Phillips. McGraw-Hill Book Company, Inc., New York, N. Y., 1947.

3. Introduction to Servomechanisms (book), A. Porter. John Wiley and Sons, Inc., New York, N. Y., 1950.

4. Reduction of Forced Error in Closed-Loop Systems, L. H. King. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 41, Aug. 1953, pp. 1037–42.

No Discussion

Some Discharge Characteristics of Lead Acid Batteries

E. A. HOXIE
MEMBER AIEE

Synopsis: This paper discusses the fundamental processes involved in the production of current in a lead acid cell, particularly as they are related to the peformance of the cell when furnishing variable or intermittent loads or a combination of both. A method of determining the size of cell required for various duty cycles is described and a general equation is derived, which, with certain restrictions, will indicate the number of positive plates required in each cell of a battery of the type selected for the application.

Behavior of Cell on Discharge

THE current produced by a lead acid storage cell when an electrically conductive circuit is established across its terminals is a result of chemical reaction of dilute sulphuric acid on the active materials, consisting of lead peroxide (PbO_2) in the positive plate and sponge lead in the negative plate.

In accordance with the generally accepted double sulphate theory, the reactions during a discharge result in the formation of lead sulphate ($PbSO_4$) at both the positive and negative plates and the acid consumed in the process is replaced by a corresponding amount of water. This is expressed by the equation

$$PbO_2+Pb+2H_2SO_4\rightarrow 2PbSO_4+2H_2O$$

As the discharge progresses, the water formed diffuses into the electrolyte to bring about the well-known reduction in specific gravity during discharge.

The internal cell voltage or the open circuit voltage, as it is commonly called, is dependent upon the concentration of acid in contact with the active materials and for a given concentration is unaffected by the discharge current. The voltage available for useful work is the voltage existing at the battery terminals during a discharge which is equal to the sum of the internal cell voltages minus the drop due to the internal resistance of the cells and the resistance of the intercell connectors. The internal cell voltage is usually within the limits of about 1.86 (with 1.050 specific gravity) to 2.14 (with 1.300 specific gravity) electrolyte at 25 degrees centigrade. There is a slight variation when grids of different alloys are used. Temperature has very little effect on the open circuit voltage, the voltage increasing with rise in temperature on the order of 0.0002 to 0.0003 volt per degree centigrade.

In observing open circuit voltage values, it should be kept in mind that the specific gravity of the free electrolyte does not necessarily indicate the concentration of acid in the pores of the plates. Also, particularly in the case of overdischarged cells or cells in poor condition, the small current taken by many voltmeters may result in readings which are lower than the correct values.

The reduction of acid concentration in the plates as the cell is discharged is also accompanied by an increase in internal resistance which increases gradually at first and then rapidly as the cell approaches full discharge, when it may be two or three times as much as for a fully charged cell. The value of the internal resistance is so small that it has little effect on the terminal voltage except when high discharge rates are encountered.

Since both the internal or open circuit voltage and the internal resistance of the cell are subject to considerable variation, it is readily seen that any attempt to use the internal resistance and the open circuit voltage of a cell as a means of estimating voltage on discharge may lead to erroneous results.

Another factor contributing to the decline in voltage during discharge is the accumulation of lead sulphate in the plates. This lead sulphate is a nonconductor and it has a tendency to clog the pores of the plates, thus impeding diffusion of electrolyte.

When the cell is discharged at low current rates, the formation of water and of lead sulphate proceeds slowly, thus allowing the higher concentration of acid in the electrolyte to readily diffuse into the pores of the plates. The voltage therefore drops slowly. However, when the cell is discharged at very high current rates, the depletion of acid in the plates takes place so rapidly that the rate of acid replacement by diffusion cannot keep pace with the production of water, thus causing a more rapid decline in voltage. This more rapid depletion of acid in the pores of the plates is the major reason for the lower ampere-hour capacities at high rates.

Paper 54-177, recommended by the AIEE Chemical, Electrochemical and Electrothermal Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted March 17, 1953; made available for printing November 30, 1953.

E. A. Hoxie is with The Electric Storage Battery Company, Philadelphia, Pa.

However, after a high rate discharge, a considerable amount of active material which has not been reduced to lead sulphate remains in the plates and additional useful acid is still present in the free electrolyte. If the cell is allowed to rest on open circuit for some time after a high rate discharge, the acid concentration in the plates is at least partially restored by diffusion, thus making additional capacity available.

As an example, it has been demonstrated that a cell may be discharged intermittently at selected high rates with periods of rest between discharges in which the total ampere-hours furnished in this manner may be upwards of 90 per cent of the ampere-hour capacity of the cell when discharged continuously at rated amperes for the same elapsed period. However, in a discharge of this kind, the discharge voltage during the intermittent charges will, of course, be lower than the discharge voltage when discharging at the lower constant rate.

Specific Gravity Readings as an Indication of State of Charge

It was stated previously that the acid consumed in the production of current is replaced by a corresponding amount of water. This water is produced at the positive plate and, since it is lighter than the free electrolyte, it has a tendency to rise to the top of the cell as it diffuses with the electrolyte. The convection currents set up by this upward movement also tend to carry the heavier acid toward the top of the cell. This movement can be readily observed in a large open tank battery. The result is that the specific gravity of the top electrolyte as measured by a hydrometer can be used as a fairly correct indication of the state of charge, providing of course, that the full charge specific gravity and the spread in specific gravity from full charge to full discharge, at some such rate as the 8-hour rate, is known.

While this paper deals with the discharge characteristics of the battery, it might be in order to state that the conditions during a charge of the battery are quite different. In this case, acid is formed in the plates and since this concentrated acid is heavier than the free electrolyte, it tends to fall to the bottom of the cell. During this time, the specific gravity of the top electrolyte may be appreciably lower than the average concentration in the cell, with the result that the true state of charge of the battery will not be indicated. However, when the rate of charge exceeds the maximum current that can be absorbed in the charging process, in which lead sulphate is reduced in both plates, the excess current produces hydrogen and oxygen gas which is forced out of the plates. These gas bubbles rise to the surface producing upward convection currents which eventually mix the acid so as to produce a concentration in the top electrolyte which indicates the true state of charge.

Effect of Temperature on Capacity

The effect of temperature on capacity of the lead acid cell is a fairly comprehensive subject which is beyond the scope of this paper. This subject will therefore be dealt with in a superficial manner.

The capacity of a cell is reduced as the cell temperature is reduced, the amount of reduction depending upon the rate of discharge and on the design of the cell. The principle reasons for the reduction in capacity are: 1. a higher viscosity of the electrolyte which tends to impede diffusion, and 2. a higher internal resistance due to an increase in the resistance of the electrolyte.

While the metallic structure of the cell has a lower resistance at reduced temperatures, the effect of this lower resistance is small as compared with the increased resistance of the electrolyte. The most important factor, particularly at high discharge rates, is the slow rate of diffusion caused by the high viscosity of the electrolyte.

Table I

Discharge Period	Per Cent Reduction in Current for each F Below 77 F
24 hours	0.52
12 hours	0.54
8 hours	0.56
3 hours	0.60
1 hour	0.64
1 minute	0.8

It is common practice to refer to the capacity of a cell at low temperature as a percentage of its capacity at a normal temperature of 77 degrees Fahrenheit (F). By common usage this usually means that a cell which is stated to have, say 60 per cent capacity at a reduced temperature, will deliver a certain current at the reduced temperature for 60 per cent of the time it will deliver this current at 77 F.

However, there are many applications in which the required discharge time is fixed and in these cases it is more convenient to refer to the capacity at low temperature as the percentage of current that can be delivered at the low temperature for the same time as at 77 F. An example of such an application is a battery used to furnish emergency light and power for a certain protection period. Incidentally, the percentage of capacity at low temperature expressed as a percentage of discharge time is lower than the percentage of capacity expressed in current for the same time. The reduction in capacity due to low temperature

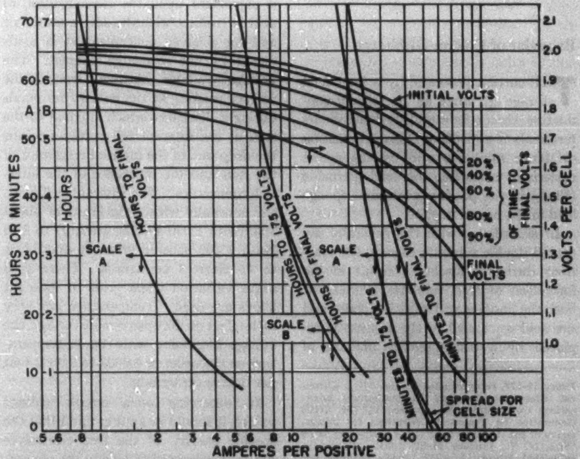


Fig. 1. Rated discharge characteristics of a typical glass jar cell. 1.200–1.220 specific gravity, 77 F

Table II

Duration of Discharge	Amperes per Positive Plate to 1.75 Volts per Cell
8 hours	5
7 hours	5.6
6 hours	6.3
5 hours	7.2
4 hours	8.5
3 hours	10.3
2 hours	13.4
1 hour	20.0
45 minutes	23.1
30 minutes	27.7
15 minutes	35.5

is greater as the discharge rate is increased. The rate of capacity reduction also increases as the temperature is lowered.

While the reduction in capacity is not linear, the capacity reduction over the range of 77 to 0 F is sometimes considered, for practical purposes, as a straight line function. The values given in Table I indicating the percentage reduction in available current from 77 to 0 F for various discharge times to 1.75 volts per cell have been used as applying to glass jar cells of 1.210 full charge specific gravity.

Selection of Type and Capacity of Battery

A storage battery is most commonly rated on the basis of ampere-hours or, to be more specific, the capacity is usually stated in terms of amperes that the cell will deliver for a given time at a specific temperature to a specified voltage such as 1.75 volts per cell, or to the knee of the volt-time curve, commonly referred to as the "final" voltage. Stationary glass jar cells are commonly rated to a voltage of 1.75 volts per cell as well as to the final voltage.

The capacity of cells containing the same components but having a varying number of plates is proportional to the number of positive plates. In conventional design, the number of negative plates exceeds the number of positive plates by one so that both sides of each positive plate are useful in producing capacity.

The nature of the service and the economic considerations will usually indicate that a particular type of battery is best suited for an application. After the decision has been made as to the type of battery to be used, the rated discharge characteristics published by the manufacturer should be consulted.

Fig. 1 shows one form of presenting discharge characteristic data for a particular type of cell which may be considered as a typical glass jar cell in 1.200–1.220 electrolyte. With this form of curve, it is possible to read the ampere capacity of a positive plate for any time from 1 minute to 72 hours to both 1.75 and final volts per cell and also to estimate the capacity over the same wide range of time to any particular voltage within the limits of the initial and final voltages.

If the requirement covers a continuous discharge at a fixed current rate, it is a simple matter to determine the size of the battery in terms of positive plates required in each cell. However, if the duty cycle consists of discharges varying in amounts, or if the discharges occur intermittently, the determination of the required capacity is more involved.

In the following examples, it has been assumed that the type of battery selected is the typical glass jar type whose discharge characteristics are shown on Fig. 1. It has also been assumed that 1.75 volts per cell is the limiting voltage and the operating temperature is 77 F. For convenience, a list of ratings taken from Fig. 1 are given in Table II.

The method of determining the size of cells required for various duty cycles, which is described in the following, has been in use for some time, but so far as is known, its use has been confined to a small group closely associated with the application of storage batteries. From time to time, its accuracy has been checked by tests which have demonstrated its reliability.

FIG. 2: ANALYSIS OF CONSTANT RATE DISCHARGE

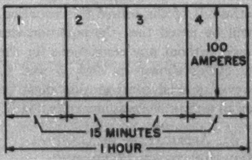


Fig. 2

A logical approach to an understanding of the procedure described in the following is to consider first the requirements of a simple 1-hour discharge at constant current as diagrammed in Fig. 2. Obviously this requirement of 100 amperes for 1 hour requires 5 positive plates which is obtained by dividing 100 by 20, 20 amperes being the rating per positive plate for 1 hour to 1.75 volts per cell.

For this 1-hour period, we may consider that the cell capacity, expressed in terms of positive plates, required for the first 15 minutes of the 1-hour discharge (period 1) is equal to the capacity required for 1 hour minus the capacity required for the remaining 45 minutes. Also, in a similar manner, the capacity required for the second 15 minutes (period 2) is equal to the capacity required to deliver 100 amperes for 45 minutes minus the capacity required to deliver 100 amperes for the remaining 30 minutes. The capacity required for the third 15-minute period (period 3) is equal to the capacity required to deliver 100 amperes for 30 minutes minus the capacity required to deliver 100 amperes for the remaining 15 minutes, and for the final 15 minutes (period 4) sufficient capacity must be provided to supply 100 amperes for 15 minutes. The sum of these individual requirements equals the number of positive plates required for the entire period of 1 hour.

The ratings involved in the determination of the number of positive plates required for the entire 1-hour period and for the four 15-minute periods of Fig. 2 are the ratings in amperes per positive plate for 15 minutes, 30 minutes, 45 minutes, and 1 hour, which for convenience are listed below:

15 minutes: 35.5 amperes per positive plate
30 minutes: 27.7 amperes per positive plate
45 minutes: 23.1 amperes per positive plate
60 minutes: 20.0 amperes per positive plate

Period 1: $\frac{100}{20}-\frac{100}{23.1}=0.67$ positive plate

Period 2: $\frac{100}{23.1}-\frac{100}{27.7}=0.72$ positive plate

Period 3: $\frac{100}{27.7}-\frac{100}{35.5}=0.80$ positive plate

Period 4: $\frac{100}{35.5}$ $=2.81$ positive plates

Total 5.00 positive plates

It may be noted that the first 15 minutes of the 1-hour discharge require only 0.67 positive plate, although a 15-minute discharge by itself would require 100/35.5 or 2.81 positive plates. Considerably more capacity is required for the last part of the discharge than for the first part. It may be stated that the conditions are most favorable for the production of current at the start of the discharge.

FIG. 3: ANALYSIS OF INCREASING RATE OF DISCHARGE CYCLE

Consider now the duty cycle illustrated in Fig. 3. The requirements of

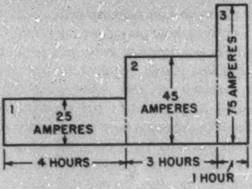


Fig. 3

periods 1, 2, and 3 may be determined by using the same method used in determining the requirements of periods 1, 2, 3, and 4 of Fig. 2.

For period 1 of Fig. 3 the number of positive plates required equals the number to deliver 25 amperes for 8 hours minus the number to deliver 25 amperes for the remaining time of 4 hours. Also, the number of positive plates needed for period 2 equals the number required to deliver 45 amperes for 4 hours minus the number required to deliver 45 amperes for 1 hour.

After again consulting the tabulation of capacities per positive plate for the times involved of 8 hours, 4 hours, and 1 hour, we find that the requirement for each period is as follows:

1 hour: 20.0 amperes per positive plate
4 hours: 8.5 amperes per positive plate
8 hours: 5.0 amperes per positive plate

Figuring the requirements for periods 1, 2, and 3 as was done in Fig. 2:

Period 1: $\frac{25}{5.0}-\frac{25}{8.5}$

Period 2: $\frac{45}{8.5}-\frac{45}{20}$

Period 3: $\frac{75}{20}$

Combining these in a single statement in which P = the total number of positive plates required we have

$$P=\frac{25}{5.0}-\frac{25}{8.5}+\frac{45}{8.5}-\frac{45}{20}+\frac{75}{20}$$

which may be written

$$P=\frac{25}{5.0}+\frac{45-25}{8.5}+\frac{75-45}{20}=8.85 \text{ positive plates} \quad (1)$$

The three terms in equation 1 no longer refer to the requirements of periods 1, 2, and 3. The first term 25/5 represents the capacity expressed as the number of positive plates required to deliver the period 1 current of 25 amperes for the total time of the cycle which is 8 hours. The second term represents the capacity required to furnish that portion of the period 2 current that is in excess of the current for period 1 for the combined time of periods 2 and 3, and the third term represents the capacity required to deliver the period 3 current which is in excess of the current of period 2 for the time of period 3.

In other words, we are considering that the first requirement of 25 amperes is required not only for 4 hours but for 8 hours and that an additional current of 20 amperes is required for 4 hours. In this way we have taken advantage of the larger capacity of the cell when the discharge lasts a longer time.

FIG. 4: CONVERSION OF ANALYSES TO A GENERAL EQUATION

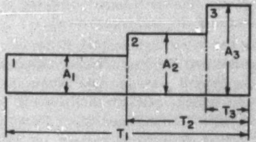


Fig. 4

Fig. 4 is shown to assist in converting equation 1 into a general equation which may be stated as

$$P=\frac{A_1}{R_1}+\frac{A_2-A_1}{R_2}+\frac{A_3-A_2}{R_3}+\ldots\frac{A_n-A_{n-1}}{R_n} \quad (2)$$

in which

P = number of positive plates required
$A_1, A_2, A_3 \ldots$ = amperes for periods 1, 2, 3 . . .
$T_1, T_2, T_3 \ldots$ = time in hours as indicated
$R_1, R_2, R_3 \ldots$ = rate per positive plate in amperes for times $T_1, T_2, T_3 \ldots$ respectively.

The following examples are given as illustrations of the use of this equation. It will be noted that the equation can be used without any restrictions for the duty cycles shown in Figs. 5 and 6. However, special consideration must be given to a duty cycle as shown in Fig. 7.

FIG. 5: APPLICATION OF GENERAL EQUATION TO AN INCREASING RATE OF DISCHARGE CYCLE

For this type of cycle, or in any cycle in which the current required for each period is higher than for the preceding period, the equation will apply without any restrictions. We have a capacity per positive plate of

1 hour: 20.0 amperes
4 hours: 8.5 amperes

$$P=\frac{25}{8.5}+\frac{50-25}{20}=4.19 \text{ positive plates}$$

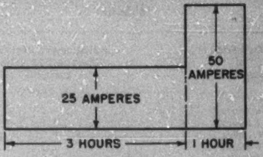


Fig. 5

Since it is necessary to use five positive plates, which is more than the theoretical calculated value of 4.19, a question may be asked as to how many amperes can actually be supplied for the final hour after supplying 25 amperes for 3 hours. It is a simple matter to apply the equation to obtain this value since the value of P is known. In other words, we have

$$5=\frac{25}{8.5}+\frac{X-25}{20}$$

Solving for X we find that, if five positive plates are used, 65 amperes can actually be furnished for the last hour.

However, if the question is asked as to how long 50 amperes may be furnished after supplying 25 amperes for 3 hours, we find that if we substitute values in the equation we have two unknowns. Since we do not know the time of the last period, we also do not know the time for the entire duty cycle and therefore we do not know the value of either discharge rate. We therefore have

$$5=\frac{25}{X}+\frac{50-25}{Y}$$

It is therefore necessary to resort to cut-and-try methods in this case.

FIG. 6: APPLICATION OF GENERAL EQUATION TO A DECREASING RATE OF DISCHARGE CYCLE

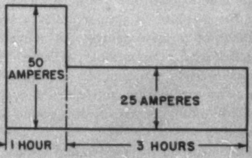


Fig. 6

It will be noted that the currents required and the duration of these currents are the same as in Fig. 5, but in this case the higher current occurs first. For this duty cycle, we have the following positive plate capacities

3 hours: 10.3 amperes
4 hours: 8.5 amperes

$$P = \frac{50}{8.5} + \frac{25-50}{10.3} = 3.45 \text{ positive plates}$$

Note that while the total ampere-hour discharge is the same for both Figs. 5 and 6, only 3.45 positive plates are required for the cycle shown in Fig. 6 while 4.19 positive plates are required for Fig. 5. This should be kept in mind in cases where a high current demand may be superimposed on a base load at any time. To be safe, it should always be assumed that the high current demand will occur during the closing minutes or hours of the discharge cycle.

For a duty cycle of this type, in which the current for the last period is less than for the preceding period, the answer to the question of how long can 25 amperes be supplied after a battery containing 4 positive plates has supplied 50 amperes for 1 hour can be answered as follows.

For this type of duty cycle, the total capacity of the cell in ampere-hours is determined by the discharge rate for the last period, which in this case is 25 amperes. Considering that the cell contains four positive plates, the discharge per positive plate at 25 amperes is 25/4 or 6.25 amperes. By consulting the rating curves in Fig. 1, we find that one positive plate will produce 6.25 amperes for 6.1 hours or, in other words, one positive plate has an ampere-hour capacity of 6.1×6.25 or 38.2 ampere hours when discharging at this rate. Since there are four positive plates in the cell, the cell has an ampere-hour capacity of 4×38.2 or 152.8 ampere-hours for a discharge at 25 amperes. Deducting the 50 ampere-hours taken out during the first hour, we have 102.8 ampere-hours remaining, which is available if the battery is discharged at 25 amperes. The cell is therefore capable of delivering 25 amperes for 102.8/25 or 4.11 hours, after delivering 50 amperes for 1 hour.

FIG. 7: RESTRICTION IN THE USE OF THE GENERAL EQUATION

This example has been included to illustrate the restriction to the use of general equation 2 as previously mentioned. In this case, the following ratings per positive plate are involved:

1 hour: 20.0 amperes
3 hours: 10.3 amperes
4 hours: 8.5 amperes

$$P = \frac{100}{8.5} + \frac{15-100}{10.3} = 3.52 \text{ positive plates}$$

While the use of the equation indicates that 3.52 positive plates are required, it is necessary in a cycle of this kind in which a peak demand precedes a smaller current demand to determine also the capacity required to meet the peak demand by itself. In this case, 100 amperes are required for 1 hour and since the 1-hour rate per positive plate is 20 amperes, it is found that 5 positive plates are required to furnish 100 amperes for 1 hour. Therefore, the cell must contain five positive plates.

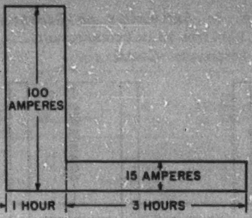


Fig. 7

The use of five positive plates means that the cell is completely discharged as far as its 1-hour capacity is concerned at the end of 1 hour, but additional capacity is available at a lower rate. In this case, it is found that while it is necessary to use five positive plates to meet the duty cycle as diagrammed in Fig. 7, actually, when five positive plates are used, the battery is capable of supplying 15 amperes for 8.3 hours after it has supplied 100 amperes for 1 hour. It is also found that a cell containing five positive plates is capable of supplying 30.3 amperes for 3 hours following a discharge of 100 amperes for 1 hour. The method of determining these values of current and time is as described for Figs. 5 and 6.

This restriction on the use of the application of general equation 2 may be stated as follows:

Restriction Concerning Use of General Equation 2

When determining the size of cell needed to meet the requirements of a duty cycle in which the current required for any period is less than for the preceding period, it is necessary to check the requirement of that portion of the cycle up to and including the period in which the high rate occurs and to compare the required capacity with the capacity required for the complete duty cycle. The required size of cell will be determined by the larger of the two requirements.

It may be noted that in the case of Fig. 6 the requirement for the first high current period of 50 amperes for 1 hour is 50/20 or 2.5 positive plates which is less than the value of 4.23 obtained by the use of the equation. The value obtained by the use of the equation is the actual requirement in this case.

FIG. 8: APPLICATION OF GENERAL EQUATION TO A MORE INVOLVED DUTY CYCLE

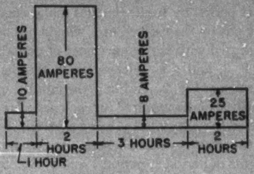


Fig. 8

This example is included to illustrate the application of the equation to a more complex duty cycle. In order to be sure of the number of positive plates required for a complicated duty cycle, the safe procedure is to consider the requirements of the duty cycle in sections starting with the first period in itself, and then the first and second period by themselves, followed by the first three periods by themselves, etc., until the entire duty cycle is covered.

In this particular case, it is fairly obvious that the large requirement of the second period (80 amperes for 2 hours) may determine the size of battery required, but in order to illustrate the method just referred to, the complete procedure has been followed. The following capacities per positive plate are involved:

1 hour: 20.0 amperes
2 hours: 13.4 amperes
3 hours: 10.3 amperes
5 hours: 7.2 amperes
6 hours: 6.3 amperes
7 hours: 5.6 amperes
8 hours: 5.0 amperes

For this cycle general equation 2 is

$$P = \frac{A_1}{R_1} + \frac{A_2 - A_1}{R_2} + \frac{A_3 - A_2}{R_3} + \frac{A_4 - A_3}{R_4}$$

Applying the equation to the complete cycle

$$\frac{10}{5} + \frac{80-10}{5.6} + \frac{8-80}{7.2} + \frac{25-8}{13.4} = 5.77 \text{ positive plates}$$

However, the first period only requires

$$\frac{10}{20} = 0.5 \text{ positive plates}$$

The first two periods only require

$$\frac{10}{10.3} + \frac{80-10}{13.4} = 6.19 \text{ positive plates}$$

The first 3 periods only require

$$\frac{10}{6.3}+\frac{80-10}{7.2}+\frac{8-80}{10.3}=4.31 \text{ positive plates}$$

The maximum requirement of 6.19 positive plates determines the size of cell. Seven plates or a total of 15 plates are required.

The capacity as just indicated for the first period by itself, the first and second periods by themselves, etc., should not be confused with the requirement of periods 1, 2, 3, and 4 as referred to in the discussion of Fig. 2. In Fig. 2 the number of positive plates required for each period referred to each period as being a part of the entire duty cycle, whereas the requirements of period 1, period 1 and 2, etc., in Fig. 8 refer to the number of positive plates required for only a portion of the complete duty cycle without regard to the balance.

FIG. 9: APPLICATION OF GENERAL EQUATION TO AN INTERMITTENT DISCHARGE CYCLE

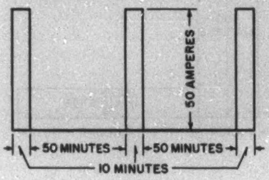


Fig. 9

This duty cycle is included to illustrate the application of the equation to cycles made up of intermittent current demands. It will be noted that the periods of open circuit are treated as a period of the complete discharge cycle. In this case, the following capacities per positive plate are involved:

10 minutes: 39.5 amperes
60 minutes: 20.0 amperes
70 minutes: 18.5 amperes
120 minutes: 13.4 amperes
130 minutes: 12.75 amperes

Applying the equation

$$\frac{50}{12.75}+\frac{0-50}{13.4}+\frac{50}{18.5}+\frac{0-50}{20}+\frac{50}{39.5}$$

$$=1.66 \text{ positive plates}$$

References

1. OPERATING CHARACTERISTICS OF LEAD-ACID STORAGE BATTERIES. *Bulletin 209*, The Electric Storage Battery Company, Philadelphia, Pa.

2. EXIDE BATTERIES FOR SWITCHGEAR CONTROL AND EMERGENCY LIGHT AND POWER. *Bulletin 210*, The Electric Storage Battery Company, Philadelphia, Pa., pp. 5–8.

3. STORAGE BATTERIES (book), George Wood Vinal. John Wiley and Sons, Inc., New York, N. Y., 3rd ed., 1940, chap. 5.

No Discussion

Some Aspects of the Charge and Discharge Processes in Lead-Acid Storage Batteries

D. NORMAN CRAIG
NONMEMBER AIEE

WALTER J. HAMER
MEMBER AIEE

THE equilibrium relations of lead-acid storage batteries have been extensively studied. The double-sulfate theory proposed by Gladstone and Tribe[1] has been substantiated by analytical methods,[2] thermodynamic studies,[3] and electromotive force measurements of galvanic cells simulating the charged state of the electrochemical system.[4] It is well recognized that on discharge one equivalent each of lead dioxide and lead and two equivalents of sulfuric acid are consumed and that two equivalents of lead sulfate and two equivalents of water are formed per faraday. It is known that the reverse occurs on charge and that the lead-acid storage battery may be subjected to many cycles of charge and discharge. The battery is reversible in that chemical and electric energy may be interconverted in repeated cycles. In practice, reversibility in the thermodynamic sense is not completely realized and the voltage during discharge is somewhat lower than the reversible electromotive force. Likewise the charging voltage is correspondingly higher. The more rapid the charging and discharging the greater is the deviation from thermodynamic reversibility and the greater is the loss in energy efficiency.

It is also known that the electromotive force of the cell increases with acid concentration. It is known that the heat of the reaction of the lead-acid accumulator based on calorimetric data agrees with the heat of reaction as calculated from the electromotive force and the electromotive-force–temperature coefficient of the cell. All in all, therefore, the relations of lead-acid storage cells under equilibrium conditions are well established and well understood.

It is also well known that the output in ampere-hours is a direct function of specific gravity or concentration of the battery acid,[5,6] that the output depends on the rate of charge and discharge, that the rate of diffusion of acid into and out of the plate pores affects the performance, and that less output is obtained as the battery ages. However, little has been published on the intrinsic characteristics of the charge and discharge processes with changing concentrations of hydrogen and sulfate ions.

It is the purpose of this paper to give some detailed information on the role the electrolyte plays in the charge and discharge processes of lead-acid batteries. The characteristics of both badly sulfated negative plates and fully developed plates will be considered. Before proceeding with the presentation of these results, we shall discuss briefly what changes occur in the specific gravity of the electrolyte of a lead-acid battery during the charging process.

Paper 54-180, recommended by the AIEE Chemical, Electrochemical and Electrothermal Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 24, 1953.

D. NORMAN CRAIG and WALTER J. HAMER are with the National Bureau of Standards, Washington, D. C.

Acknowledgment and appreciation are herewith expressed to C. G. Malmberg for the data on electric resistivity, to R. L. Cottington for the data on viscosity, to Bruce Bendigo for the data on solubility, to J. D. Appel for much of the data on the rate of conversion of lead sulfate to lead, and to C. L. Snyder for the observations on the battery made with sodium and magnesium sulfates, presented in this paper. Acknowledgment is also made to Dr. W. J. Youden and Dr. Churchill Eisenhart for consultation during the course of these investigations.

Table I. Specific Gravities of Electrolyte and Rate of Gassing During Charge

	Cell 1 Charge at 6 Amperes			Cell 2 Charge at 10 Amperes		
Charge, Ampere-Hours	Specific Gravity (Sp. Gr.) of Electrolyte	Δ Sp. Gr.	Gas, Millimeters per Minute	Sp. Gr. of Electrolyte	Δ Sp. Gr.	Gas, Millimeters per Minute
0	1.152		0.0	1.125		0.0
15	1.150	0.007	0.0	1.128	0.003	0.1
30	1.173	0.014	0.2	1.137	0.009	0.4
45	1.191	0.018	0.5	1.154	0.017	0.6
60	1.212	0.021	2.0	1.180	0.026	2.8
75	1.265	0.053	32.0	1.260	0.080	50.0
90	1.275	0.010	54.0	1.276	0.016	98.0

Table II. Specific Gravities of Electrolyte at Bottom and Top of Two Cells During Charge at 10-Ampere Rate

	Cell 1			Cell 2		
Charge, Ampere-Hours	Sp. Gr. Bottom	Sp. Gr. Top	Sp. Gr. Difference	Sp. Gr. Bottom	Sp. Gr. Top	Sp. Gr. Difference
0	1.103	1.103	0.000	1.115	1.115	0.000
15	1.137	1.115	.022	1.152	1.130	.022
30	1.169	1.131	.038	1.185	1.147	.038
45	1.200	1.150	.050	1.217	1.167	.050
60	1.227	1.172	.055	1.242	1.189	.053
75	1.250	1.201	.049	1.264	1.216	.048
90	1.267	1.244	.023	1.279	1.265	.014
100	1.275	1.268	.007	1.285	1.280	.005

Changes in Specific Gravity During Charge

It is known that transient concentration gradients exist within a battery on charge or discharge. These gradients produced on charge frequently lead the user to the conclusion that the battery is not taking a charge. When a discharged battery is first put on charge the heavy sulfuric acid formed by the charging process (reaction 3), tends to settle to the bottom of the battery. Hence, the usual test of the top liquid with a hydrometer will show at first only a small rise in specific gravity and it may be falsely concluded that the battery "is not taking a charge." With further charging, diffusion and the mixing caused by the bubbles evolved when gassing is reached bring the dense solution to the top where it mixes with the less dense solution and the specific gravity of the solution at the top of the plates than rises rapidly and the hydrometer reading becomes significant. If a salt, for instance, were added between the two readings, one might falsely conclude that the salt had made possible the charging of a battery that "would not take a charge." Also, on standing, the dense solution at the bottom of the battery formed on initial charge will diffuse into the less dense solution until uniformity in acid composition throughout the battery is attained. Likewise, if a salt were added and sufficient time were allowed for diffusion of the dense solution into the less dense solution, one might also falsely conclude that the salt had charged the battery or had brought the specific gravity up by more than could be accounted for by its mere addition.

Although these effects are probably well recognized it seems worth while to give additional data. We shall consider first the changes that occur in the specific gravity of the electrolyte during charge when measured in the usual way, that is, at the top of the cell, and the changes that occur when the rate of gassing becomes appreciable. Data obtained on two automotive cells at charging rates of 6 and 10 amperes are given in Table I. The table shows the specific gravity of the electrolyte at the top of the two cells and the rate of gassing for successive increments of charge. As the charge progresses it is observed that the rise in specific gravity increases with successive increments of charge and that the most rapid rise occurs when the increase in the rate of gassing is greatest whereby the mixing of the electrolyte is enhanced.

Next, we shall present a more detailed study of the changes in the specific gravity of the electrolyte during charge. In order to measure the specific gravity of the electrolyte at the bottom and top of the cells during charge, some cells of the automotive type were provided with valves at the bottom of the cells whereby electrolyte could be periodically removed from the bottom for hydrometer readings. The hydrometer readings of the electrolyte at the top of the cells were taken in the usual manner. To eliminate any initial concentration gradients, the electrolyte was drained from each cell and thoroughly mixed. It was then returned to its respective cell and this process was repeated until the specific gravity of the electrolyte at the top and bottom of the cell was the same.

In Table II the specific gravity at the bottom and at the top of each of two cells differing only slightly in original specific gravity are shown during the charge. The differences in specific gravities for top and bottom are also listed. It is seen that these differences are quite uniform for any given increment of charge and that the two specific gravities (top and bottom) again approach a common value at the end of charge when gassing, which then becomes pronounced, mixes the lower and top portions of the electrolyte in each cell.

To ascertain the magnitude of the foregoing phenomena on the estimation of state of charge as determined by hydrometer readings, two cells were charged in series; the electrolyte in one cell was continuously circulated and the electrolyte in the other cell was not circulated. Data are given in Table III. The results for the first part of the charge show that the increase in the specific gravity of the electrolyte for a given increment of charge is less for the electrolyte at the top of the

Table III. Comparison of Specific Gravities of Electrolyte in Top of Cell with Specific Gravities Observed for Circulated Electrolyte. Charged at 10 Amperes

	Cell 1, Uncirculated Electrolyte			Cell 2, Circulated Electrolyte		
Charge, Ampere-Hours	Sp. Gr.	Δ Sp. Gr.	Calc.,* Per-Cent Charge	Sp. Gr.	Δ Sp. Gr.	Calc., Per-Cent Charge
0	1.102			1.103		0
15	1.115	0.013	7.8	1.130	0.027	17.2
30	1.131	0.016	17.5	1.158	0.028	35.0
45	1.150	0.019	28.9	1.185	0.027	52.2
60	1.172	0.022	42.2	1.210	0.025	66.2
75	1.201	0.029	59.6	1.232	0.022	82.2
90	1.244	0.043	85.5	1.254	0.022	98.1
100	1.268	0.024	100.0	1.260	0.006	100.0

* Apparent.

Fig. 1. Equipment for charging and discharging negative plates

cell in which the electrolyte is not circulated than for the increase observed for the cell in which the electrolyte is circulated. During the first part of the charge the charging efficiency is high and the linear relationship observed for the circulated electrolyte is in accord with theory. During the latter part of the charging period when gassing ensues the charging efficiency is lowered; nevertheless the rate of the rise in specific gravity of the electrolyte at the top of the cell containing the uncirculated electrolyte increases sharply because the gassing causes effective mixing of the more dense electrolyte in the lower part of the cell with that above the plates.

If we consider that a linear relation exists between state of charge and the specific gravity we can calculate the per cent of charge at each stage of the charging process. These percentages are listed in Table III. It will be noted that after 3 hours of charge (30 ampere-hour input) the specific gravity of the uncirculated cell indicates that the cell had received only 17.5 per cent (third column) of charge, whereas in fact it had received about 35 per cent (see last column) of full charge. If now a salt were added to the uncirculated cell and sufficient time allowed, the specific gravity in cell 1 (uncirculated) would rise to about 1.158 (disregarding any rise in specific gravity due to addition of the salt) and one might falsely conclude that the salt had charged the cell by about 17 per cent.

Chemical Reactions in Lead-Acid Batteries During Charge

The reactions that occur at the negative and positive plates in lead-acid storage batteries during charge may be formulated as follows:

Negative plate:

$$PbSO_4+2\epsilon\rightarrow Pb+SO_4^{--} \quad (1)$$

Positive plate:

$$PbSO_4+2H_2O\rightarrow PbO_2+4H^{+}+SO_4^{--}+2\epsilon \quad (2)$$

Over-all cell reaction during charge:

$$2PbSO_4+2H_2O\rightarrow Pb+PbO_2+2H_2SO_4 \quad (3)$$

where ϵ is the electron and the other symbols have their usual significance. The plate potentials during charge can be measured relative to a reference cadmium electrode, and the algebraic difference in these potentials is equal to the measured cell voltage. The charging efficiency or the rate of conversion of lead sulfate to lead and lead dioxide may be determined electrically by discharge or chemically by analysis of lead and lead dioxide and/or sulfuric acid produced during the charge. In practice only the former is done.

In order to study the charging efficiencies of pasted plates (negative and positive) separately, each plate, negative or positive, can be charged against pure lead or unformed Planté plates as dummies. The latter have negligible capacity because of their nonporous surface and do not affect the amount of acid formed during charge nor the amount consumed during discharge. The early literature records cases where 2 years were required to form Planté plates for usable service.

When a pasted negative plate is charged against unformed Planté plates as positives the following reactions occur:

Negative pasted plate:

$$PbSO_4+2\epsilon\rightarrow Pb+SO_4^{--} \quad (4)$$

$$2H^{+}+2\epsilon\rightarrow H_2 \text{ (competing reaction)} \quad (4A)$$

(This competing reaction will be considered later in the paper.)

Unformed Planté positive plate:

$$1/2Pb+H_2O\rightarrow 1/2PbO_2+2H^{+}+2\epsilon \text{ (momentary reaction)} \quad (5)$$

$$H_2O\rightarrow 1/2O_2+2H^{+}+2\epsilon \text{ (lasting reaction)} \quad (6)$$

Over-all charge reaction (equations 4 plus 6):

$$PbSO_4+H_2O\rightarrow Pb+H_2SO_4+1/2O_2 \quad (7)$$

On charge the Planté positive becomes

Fig. 2. Detail of equipment for charging and discharging negative plates

Table IV. Effect of Acid Concentration on Rate of Conversion of Lead Sulfate of Hard Sulfated Negative Plates to Lead and Sulfate Ion. Charging Rate 1 Ampere, Discharging Rate 2 Amperes

Cycle	Input, Ampere-Minutes	Sp. Gr. 1.021	Sp. Gr. 1.099	Sp. Gr. 1.200	Sp. Gr. 1.301
		Grams H_2SO_4/10-Milliliter Solution			
0	0.0	0.317	1.568	3.336	5.204
2	446.4	0.648	1.875	3.622	5.354
3	735.0	1.186	2.331	3.999	5.568
4	1105.0	1.728	2.845	4.399	5.832
5	1155.0	1.973	3.094	4.613	6.046
		Increments			
0–2		0.331	0.307	0.286	0.150
2–3		0.538	0.456	0.377	0.214
3–4		0.542	0.514	0.400	0.264
4–5		0.245	0.249	0.214	0.214
0–5		1.656	1.526	1.277	0.842
		Minutes of Discharge in 1.200 Sp. Gr. H_2SO_4 Solutions			
		664.7	524.7	393.5	197.7

covered immediately with a coating of PbO_2, and since the plate is not porous, penetration of the plate by electrolyte cannot ensue, and water is therefore decomposed at the plate surface and oxygen is evolved. We therefore have a method for studying the rate of conversion of lead sulfate to lead ("active material") at the negative plate (reaction 7). The amount of sulfuric acid formed may be readily determined by titration with standard alkali, and this amount, as shown by reaction 7, is equivalent to the amount of lead formed. The relative amounts of lead formed can also be determined by discharging the plates in series at constant current in fixed quantities of sulfuric acid solution containing excess sulfuric acid against Planté dummies to a negative plate versus cadmium electrode potential of 1 volt. The amount of lead produced could also be determined by chemical analysis but this would not only be time-consuming but very difficult.

Either the determination of the sulfuric acid formed or the determination by discharge of the relative amounts of lead formed is a convenient method to study the factors involved in the charging of hard dense "sulfated" negative plates which are produced when a battery is allowed to stand idle for a considerable time especially in a discharged state under conditions of fluctuating temperatures. It is well known that negative plates sulfate at much faster rates than positive plates and that the latter are frequently said to disintegrate during attempts to restore badly sulfated negative plates. We shall therefore concentrate our attention on negative plates. The principles which will be outlined now for hard sulfated negative plates will apply also to positive plates.

Old sulfated negative plates (4.875 inches by 5.6875 inches by 0.125 inch nominal dimensions) were removed from discarded batteries, inspected for mechanical soundness, and the rate with which the lead sulfate on their surface and in the plate pores was converted to lead during the charge was determined. All plates used in these studies were tested and found to have no residual capacity; this, together with their extremely low initial charging efficiencies, attests to their sulfated nature. Also, in these relative studies all plates were paired; plates of any one cell of any one battery were paired for comparisons of variables under study. Each negative plate under study was spaced in a Lucite or Plexiglas container between two Planté dummies of pure lead joined in parallel. The dummy plates were approximately the same width and height as the negative plate. The assembly for these studies is shown in Figs. 1 and 2.

Since at least two and sometimes eight plates were cycled in series, each cell was provided with a separate reference cadmium electrode so that the potential of the negative plate in each cell could be quickly measured. Suitable switches made it possible to terminate the discharge of any particular plate with only a momentary interruption of the discharge current. In all except the final discharge of each plate, the discharge was terminated when the negative plate versus the cadmium electrode rose to 2.60 volts. This potential marks the end of the abrupt and characteristic rise at the end of discharge. The exact time was recorded when discharge of each plate was terminated; this time, as well as the rate of discharge and the length and rate of charge, was recorded so that the excess

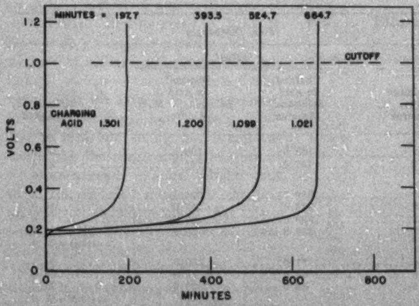


Fig. 3. Cadmium potentials and outputs of negative plates discharged in sulfuric acid solutions of 1.200 sp. gr. The plates were previously charged in sulfuric acid solutions of indicated gravities

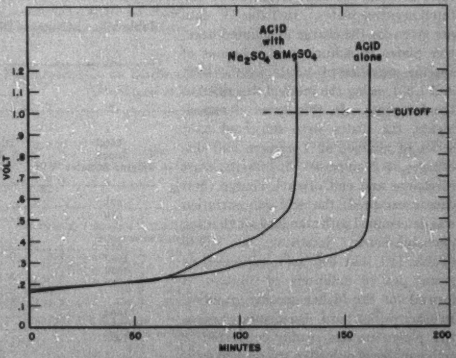


Fig. 4. Cadmium potentials and outputs of negative plates discharged in sulfuric acid solutions of 1.200 sp. gr. The plates were previously charged in the indicated solutions

Table V. Charging Efficiency for Paired Negative Plates in H_2SO_4 Solutions and H_2SO_4-Salt Solutions of Various Specific Gravities at Low Charging Rates

Sp. Gr. of Initial Acid Solutions	Excess Input, Ampere-Minutes		Rate of Charging Amperes	Acid Formed, Grams per 360 Milliliter Solution		Per-Cent Conversion		Differ-ences
	Acid	Acid + M*		Acid	Acid + M*	Acid	Acid + M*	
1.020	481.6	476.6	1	13.33	11.34	90.8	78.0	12.8
1.020	473.2	451.2	1	12.36	10.65	85.7	77.4	8.3
1.020	837.9	855.9	0.5–6†	21.09	20.23	82.5	77.5	5.0
1.020	849.3	807.9	0.5–6†	21.38	19.47	82.5	79.0	3.5
1.020	825.1	854.5	0.5–6†	21.27	18.29	84.5	70.2	14.3
1.020	872.7	869.7	0.5–6†	20.88	16.77	78.5	63.2	15.3
1.020	802.4	707.8	1	21.64	16.49	88.4	76.4	12.0
1.020	1185.6	1113.6	1	33.95	29.41	93.9	86.6	7.3
1.020	1977.0	1854.6	1	63.36	60.01	86.5	86.7	−0.2
1.020	1945.8	1868.4	1	61.70	59.18	85.1	85.0	0.1
1.020	1852.4	1675.0	1	61.34	55.84	88.7	88.3	0.4
1.100	687.6	685.6	1	16.71	15.92	79.7	76.1	3.6
1.150	705.2	707.8	1	16.67	13.76	77.5	63.8	13.7

* Mixture of sodium and magnesium sulfates.

† The input of the 1st, 3rd, and 4th cycles was at 0.5 ampere. The input of the 2nd cycle was at 6 amperes.

Table VI. Charging Efficiency for Negative Plates in Sulfuric Acid of Low Specific Gravity (1.020) and at High Charging Rates

Rate of Charging, Amperes	Excess Input, Ampere-Minutes		Acid Formed, Grams per 360 Milliliter Solution		Per-Cent Conversion		Differences
	Acid	Acid + M*	Acid	Acid + M*	Acid	Acid + M*	
1†							−0.2 to +12.8
6	1047.0	1053.6	19.04	9.43	59.6	29.3	+30.3
6	1057.2	1057.2	18.12	8.03	56.2	24.9	+31.3
6	1060.8	1058.4	17.78	8.47	55.0	26.2	+28.8
6	1712.6	1725.4	23.04	13.65	44.1	25.9	+18.2
6	1752.0	1775.0	17.17	11.56	32.1	21.4	+10.7

* Mixture of sodium and magnesium sulfates.

† See Table V.

input could be calculated for any stage in the cycling process for each plate. In the final discharge of each plate the output was calculated to a negative plate versus cadmium potential of 1 volt.

We shall consider first the effect of solutions of various concentrations of sulfuric acid on the redevelopment of sulfated negative plates. In Table IV, data are given for the charge of sulfated negative plates in sulfuric acid solutions of specific gravities of 1.021, 1.099, 1.200, and 1.301, using the method described in the foregoing. In this series of experiments, the plates were subjected to a series of charges at 1 ampere and discharges at 2 amperes. Before the start of charge and end of each charge (first charge excepted) the acid concentration was determined with standard alkali using phenolphthalein indicator. We note (Table IV) that less acid, expressed as grams per 10 milliliters of solution, is formed for the higher specific gravities. In other words, hard dense sulfated negative plates are less readily converted to active material or lead in strong acid solutions than in weak ones.

As a confirmation of these results the charged plates after the final charge (cycle 5) were removed from their respective solutions, thoroughly washed with distilled water, and discharged at 2 amperes in 1.200 specific gravity acid. The discharge curves are shown in Fig. 3 and the minutes of discharge on a 2-ampere drain to a cadmium potential of 1 volt are given at the bottom of Table IV. It is apparent, and most significantly so, that more minutes of discharge are obtained for the plates which were charged in the more dilute acid.

The question then arises as to the cause of this. Is the decreased efficiency for the conversion of hard dense lead sulfate to lead due only to the large increase in hydrogen-ion content, or in part to the increase in sulfate-ion content, or to changes in the solubility of lead sulfate with acid concentration? It is known that the solubility of lead sulfate is not a direct function of acid concentration,[7,8] and therefore cannot in itself account for the observation in the preceding paragraph. (More is presented on solubility in a later section of this paper.)

In this paper we are concerned with the effects caused by altering the sulfate-ion concentration without appreciably changing the hydrogen-ion concentration. This may be done readily by the simple expedient of adding a metallic sulfate. Metallic sulfates will alter the dissociation of sulfuric acid since sulfuric acid is known to ionize stepwise with the second step being incomplete.[9] However, this alteration of the ionization of the acid will not alter the concentration of hydrogen ion appreciably and what alteration does occur will be insignificant compared to the large changes described previously for pure aqueous sulfuric acid solutions. For this purpose, we cannot add sulfates like copper or silver sulfates as these will lead to plating of the metals (copper or silver) on the negative plates giving rise to increased self-discharge of the battery. We can, however, add sulfates of metals such as lithium, sodium, potassium, rubid-

Table VII. Minutes of Discharge in 1.200 Specific Gravity H_2SO_4 for Charged Plates Discussed in Tables V and VI

	Final Discharges			
		Output		
Total Input, Ampere-Minutes	Discharge Rate, Amperes	Charged in Acid Solutions, Minutes	Charged in Acid Solutions + M,* Minutes	Differences, Minutes
575	2	163.0	128.2	34.8
575	2	157.8	131.4	26.4
				average = 30.6
1052.5	2	238.7	224.4	14.3
1052.5	2	233.6	217.0	16.6
1052.5	2	230.5	203.5	17.0
1052.5	2	236.8	189.2	47.6
				average = 23.9
1080	6	71.0	30.5	41.5
1080	6	64.8	30.8	34.0
1080	6	63.9	28.9	35.0
				average = 36.8
1920	6	81.7	48.8	32.9
1920	6	58.3	35.8	22.5
				average = 27.7

* Mixture of sodium and magnesium sulfates.

Table VIII. Data on Solubility at 25 Degrees Centigrade of Lead Sulfate in Sulfuric Acid Solutions and Similar Solutions Containing Sodium and Magnesium Sulfates

Acid Solution, Sp. Gr.	Weight, Per Cent H_2SO_4	Micrograms $PbSO_4$ per Milliliter of Acid Solution	Sp. Gr. of 1,000-Milliliter Acid Solution Plus Salt Mixture*	Weight, Per Cent H_2SO_4	Micrograms $PbSO_4$ per Milliliter of Acid-Salt Solution
1.0050	0.789	4.64	1.022	0.768	4.89
1.0087	1.49	4.80	1.026	1.44	5.06
1.020	2.94	5.69	1.034	2.84	5.66
1.045	7.48	6.23	1.061	7.31	6.15
1.069	10.40	6.54	1.084	10.18	6.48
1.097	14.47	6.15	1.112	14.14	6.01
1.164	23.49	4.10	1.177	22.98	4.08
1.196	27.60	3.24	1.209	27.00	3.30
1.250	34.33	2.06	1.260	33.48	2.09
1.250	34.33	2.06	1.270†	32.73	2.13
1.292	39.48	1.27	1.301	38.57	1.37

* 20.0 grams $MgSO_4.7H_2O + 9.0$ grams Na_2SO_4.

† Double the amount of salt mixture.

ium, cesium, or magnesium, all of which are readily soluble in battery electrolyte. For the present purpose we have chosen a mixture of sodium and magnesium sulfates (approximately 38 per cent Na_2SO_4, 47 per cent $MgSO_4$, and 15 per cent water of hydration). These salts are readily available, are inexpensive, and have historical background.

To obtain an over-all picture of the phenomena that might occur, studies were made at various charge rates and for acid solutions varying in specific gravity from 1.020 to 1.200. For these studies, unless otherwise specified, 21 grams of the salt mixture was added to 360 milliliters of sulfuric acid of the desired specific gravity. It was found that smaller amounts produced the same phenomena but to a somewhat lesser extent.

Results of these experiments for charges at 1 ampere are given in Table V. The current density of the charge was, therefore, 0.018 ampere per square inch. As was done in the studies for pure aqueous sulfuric acid solutions, the acid concentration was determined with standard alkali before start of and at the end of each charge. Results given in Table V show that less acid is produced when a mixture of sodium and magnesium sulfates is added to the electrolyte. This means that the rate of or the efficiency for the conversion of hard dense lead sulfate to lead and sulfate ion is not enhanced by the addition of a mixture of magnesium and sodium sulfates but rather is curtailed during the charging process. It should be noted that this is true for acid specific gravities of 1.020, 1.100, and 1.150.

Similar data are given in Table VI for a much higher charging rate of 6 amperes, i.e., for a current density of 0.108 ampere per square inch. Here the detrimental effect of a mixture of magnesium and sodium sulfates on the charging efficiency or redevelopment of the negative plates is very marked; the efficiency is only about 50 per cent of that obtained for the pure aqueous solutions of sulfuric acid.

As a confirmation of these results, the charged plates of the experiments just described were removed from their respective charging solutions, washed in distilled water, and discharged in 1.200 specific gravity sulfuric acid at 2 or 6 amperes. As before, a reference cadmium electrode was immersed in the solutions and the potentials of the negative plates were measured with reference to the cadmium electrode throughout the discharge. Data for these discharges are given in Table VII, and are shown graphically for the first pair of plates in Fig. 4. Without exception, for different charging rates and for different acid specific gravities, higher ampere-minute outputs were obtained for plates that had been charged in pure aqueous sulfuric acid solutions as opposed to those charged in similar solutions to which a mixture of sodium and magnesium sulfates had been added. This shows that more lead had been produced on charge as exemplified by reaction 7. In short, titrations show that more sulfuric acid and discharges show that more lead is produced during charge in solutions free of a mixture of sodium and magnesium sulfates; in other words

Table IX. Effect of Na_2SO_4-$MgSO_4$ on Discharge of Sulfated Negative Plates Charged at Low and High Rates in H_2SO_4 Solutions of Extremely Low Specific Gravity

Charge Rate, Amperes	Input, Minutes	Specific Gravity							
		Pair 1.020	1.020 + M*	Pair 1.020	1.020 + M†	Pair 1.020	1.020 + M‡	Pair 1.020	1.020 + M§
		Minutes of Discharge at 2 Amperes							
0.5	180	5.8	6.3	5.7	6.9	5.9	7.8	5.7	9.1
0.5	605	77.0	69.2	72.0	87.7	71.5	64.0	64.0	63.5
		Minutes of Discharge at 6 Amperes							
6.0	30			0.7	1.3	0.7	0.9		
6.0	60			3.0	3.9	2.5	2.8		
6.0	75			16.4	13.3	12.9	9.6		

* 3 grams of mixture of sodium and magnesium sulfates.

† 6 grams of mixture of sodium and magnesium sulfates.

‡ 9 grams of mixture of sodium and magnesium sulfates.

§ 18 grams of mixture of sodium and magnesium sulfates.

Table X. Effect of Na_2SO_4-$MgSO_4$ on Discharges of Sulfated Negative Plates Charged and Discharged at Low Rate

Input, Minutes	Specific Gravity							
	Pair 1.020	1.020 + M*	Pair 1.020	1.020 + M*	Pair 1.020	1.020 + M*	Pair 1.020	1.020 + M*
	Charged at 1 Ampere. Discharged at 2 Amperes. Minutes of Discharge							
210	19.8	33.2	19.3	30.8	26.3	47.8		
435	59.2	85.5	64.3	81.0	87.0	119.2		
785	191.0	212.5	202.0	212.5	219.0	254.0		
1,087	345.2	349.0	355.0	364.5	331.0	365.0		
1,270	497.6	492.6	498.5	503.1	474.0	508.1		
	Charged at 0.5 Ampere. Discharged at 2 Amperes. Minutes of Discharge							
60	2.5	4.5	2.2	4.8	1.5	4.3	1.8	4.3
120	7.4	13.0	7.0	12.2	5.6	13.5	6.5	12.5
360	24.2	39.7	23.7	36.5	22.5	43.2	21.6	36.8
780	81.7	109.0	75.8	100.8	71.2	113.5	71.7	97.0
1,080	171.7	187.5	166.8	180.8	160.9	189.5	163.0	176.5

* Mixture of sodium and magnesium sulfates.

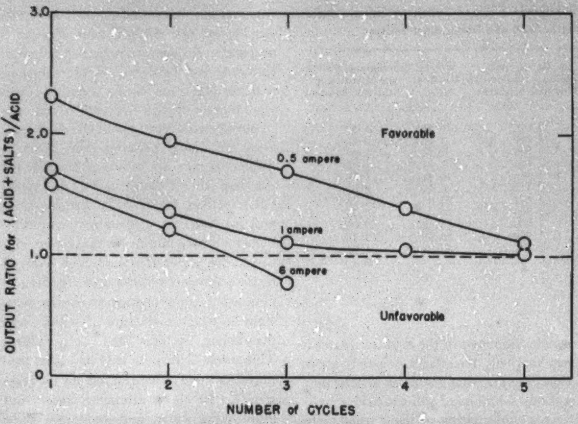


Fig. 5. The performance of negative plates in dilute sulfuric acid solutions (1.020 sp. gr.) containing various quantities of a mixture of sodium and magnesium sulfates

batteries with sulfated plates accept charge at a slower rate when sodium and magnesium sulfates are added to the electrolyte. The difference between 100-percent efficiency and that observed is due to discharge of hydrogen ions in accordance with the competing reaction 4(A); therefore, a mixture of sodium and magnesium sulfates induces more gassing or bubbling and more water losses during the charging process, especially for high-rate charges.

It will be noted from the differences in Table VII that the detrimental effect of magnesium and sodium sulfates becomes less as the plate pores are "opened-up" or when the plates are given more charge (ampere-minute input). This is true for both high and low charging rates. It is also noted that the detrimental effect is more marked the higher the charging or discharging rate, which suggests that sodium and magnesium sulfates adversely affect the diffusion of the electrolyte to and away from the plates. (More on diffusion is given in a later section of this paper.)

Therefore, a mixture of sodium and magnesium sulfates leads to decreased efficiency for conversion of lead sulfate to lead similar to that observed when the concentration of sulfuric acid is increased. It was stated before that the solubility of lead sulfate is not a direct function of sulfuric acid concentration. It remains to be determined if a mixture of sodium and magnesium sulfates alters the solubility of lead sulfate in sulfuric acid solutions. Accordingly, such solubilities were determined by the standard dithizone method[10] using a spectrophotometer. Data obtained at 25 degrees centigrade are given in Table VIII for solutions ranging in specific gravities from 1.0050 to 1.301. It is noted that the addition of the mixture of sodium and magnesium sulfates causes no significant difference in the solubility of lead sulfate in sulfuric acid solutions and, in particular, measured differences are insignificant when compared to the solubility changes as a function of specific gravity over the specific gravity range of sulfuric acid solutions encountered in normal battery operation. Therefore, the decreased efficiency in the conversion of hard dense lead sulfate to lead produced by additions of a mixture of sodium and magnesium sulfates to sulfuric acid solutions is not due to alterations in solubility of lead sulfate in sulfuric acid solutions but to the increased amount of sulfate ion resulting from the addition of mixtures of sodium and magnesium sulfates.

Chemical Reactions in Lead-Acid Batteries During Discharge

In the foregoing section it was shown that the addition of a mixture of sodium and magnesium sulfates to the electrolyte of lead-acid batteries is detrimental in the redevelopment of hard sulfated negative plates. Therefore, next we consider the effect of such a mixture on the discharging process and the relation between the charge and discharge. The over-all discharging reaction of the lead-acid storage battery is the reverse of reaction 3. The discharging reactions of the negative and positive plates are the reverse of reactions 1 and 2 respectively. For pasted negative plates against unformed Planté positive plates the reactions are:

Negative pasted plate:

$$Pb + SO_4^{--} \rightarrow PbSO_4 + 2\epsilon \qquad (8)$$

Positive unformed Planté plate:

$$1/2PbO_2 + 2H^+ + 2\epsilon \rightarrow 1/2Pb + H_2O \text{ (momentary reaction)} \qquad (9)$$

$$2H^+ + 2\epsilon \rightarrow H_2 \text{ (lasting reaction)} \qquad (10)$$

Over-all discharge reaction (equations 8 + 10):

$$Pb + H_2SO_4 \rightarrow PbSO_4 + H_2 \qquad (11)$$

For this study of the discharge process, old sulfated negative plates obtained from discarded batteries were chosen and after charge were studied on discharge in sulfuric acid solutions ranging in specific gravity from 1.020 to 1.200. In Table IX, data are given for extremely dilute solutions of sulfuric acid of specific gravity of 1.020 (this low specific gravity is rarely found in practice) for several successive cycles at low and high charge and discharge rates. At the low rates of charge (0.5 ampere) and discharge (2 amperes) a mixture of sodium and magnesium sul-

Table XI. Effect of Na_2SO_4-$MgSO_4$ on Discharge of Sulfated Negative Plates Charged a Low and High Rates in H_2SO_4 of Moderate Specific Gravity

Input, Minutes	Specific Gravity							
	Pair		Pair		Pair		Pair	
	1.098	1.098 + M*	1.098	1.098 + M*	1.098	1.098 + M*	1.098	1.098 + M*
Low: Charged at 1 Ampere. Discharged at 5 Amperes. Minutes of Discharge								
255	15.8	15.8	15.8	15.8	15.0	17.2	14.5	16.7
450	39.6	40.5	38.5	37.4	39.2	39.5	37.3	41.3
High: Charged at 6 Amperes. Discharged at 6 Amperes. Minutes of Discharge								
60			20.6	11.2	20.2	13.8		
120			55.8	42.5	56.0	48.5		

* Mixture of sodium and magnesium sulfates.

fates shows an apparent beneficial effect after a low input of 180 minutes. However, this apparent beneficial effect diminishes and becomes insignificant as the cells are cycled since less acid is formed on the charge when a mixture of sodium and magnesium sulfates is present. This is shown on the next cycle wherein the input was 605 minutes. Here the difference between the minutes of discharge obtained for the cells free of and those containing the sulfates of sodium and magnesium is insignificant (two cells with the salt mixture gave fewer minutes, one more, and one was practically equivalent to the cell containing only aqueous sulfuric acid). For the high-rate charges and discharges in this extremely dilute acid solution, the salt mixture is detrimental for the third discharge. This is shown in more detail by the data of Table X. Here again we see that a mixture of sodium and magnesium sulfates produces an apparent beneficial effect which diminishes and becomes insignificant as the cells are cycled. This phenomenon is shown graphically in Fig. 5. The initial apparent beneficial effect observed for this extremely dilute acid solution is not surprising when one considers that the sulfate content of sulfuric acid of 1.020 specific gravity is 2.85 per cent whereas the sulfate content of 360 milliliters of 1.021 specific gravity acid containing 21 grams of the mixture of sodium and magnesium sulfates is 5.97 per cent. The apparent beneficial effect disappears on cycling because less acid is produced during charge in cells to which sodium and magnesium sulfates are added, whereas cells not so treated gain more rapidly in sulfate content.

We have next to consider whether or not this effect is observed for the normal specific gravity encountered in practical operations. First, we shall consider a specific gravity of 1.100, the gravity frequently observed at the termination of discharge when the battery shows a very rapid drop in voltage. Data are given in Table XI. For the low charge rates and high discharge rates no significant difference is noted between cells with or without the salts. For high charge and discharge rates it is seen that a mixture of sodium and magnesium sulfates produces a detrimental effect; the discharge times for those containing the salt mixture are only 60 to 80 per cent of those obtained from cells containing only aqueous sulfuric acid.

Next, we shall consider results obtained for the specific gravity of 1.200, the specific gravity corresponding to about 50 per cent of full charge. Since it has already been shown that the mixture is detrimental to the charging process, the plate pairs were all charged in 1.200 acid for two cycles (input=600 ampere-minutes.) The acid was then dumped from all cells, and either (1) pure acid (specific gravity=1.200) or (2) pure acid plus a mixture of sodium and magnesium sulfates (specific gravity=1.232), or (3) acid of specific gravity (1.233) equal to that of the acid-salt mixture was added to the individual cells. Data given in Table XII show no significant differences between the discharge characteristics of the negative plates in sulfuric acid solutions of normal specific gravity and of the negative plates in similar solutions to which a mixture of sodium and magnesium sulfates was added.

Table XII. Effect of Na_2SO_4-$MgSO_4$ on Discharge of Sulfated Negative Plates Charged at Low Rate (0.5 Ampere) in 1.200 Specific Gravity Sulfuric Acid Solutions and Discharged at Low Rate (2 Amperes) in Various Solutions. Total Input=600 Ampere-Minutes

Pair		Pair		Pair		Pair	
1.200	1.200 + M*	1.200 + M*	1.200 + A†	1.200	1.200 + M*	1.200 + M*	1.200 + A†
Specific Gravities of Solutions After Above Additions							
1.200	1.235	1.232	1.233	1.200	1.235	1.232	1.233
Minutes of Discharge							
74.4	71.1	76.2	79.5	82.0	74.5	78.7	76.7

* Mixture of sodium and magnesium sulfates.

† Additional sulfuric acid.

In summary, then, we may list in Table XIII the effects observed for a mixture of sodium and magnesium sulfates on the charge and discharge of cells containing hard sulfated negative plates. In Table XIII very low, low, and moderate specific gravities refer respectively to 1.020, 1.100, and 1.200, and low and high rates refer respectively to 1.0 (or 0.5 ampere) and 6.0 amperes per plate. It will be noted that a mixture of sodium and magnesium sulfates leads to different results at low and high rates and is more detrimental at the higher rates. The beneficial effect at the very low specific gravity and low discharge rate vanishes on cycling. The differences observed between the behaviors at low and high rates might well be explained if the presence of the mixture of sodium and magnesium sulfates significantly increases the viscosity or electric resistivity of sulfuric acid solutions. Accordingly, these quantities were determined, by accepted methods, for pure aqueous sulfuric acid solutions and for acid solutions containing the mixture of sulfates.

Table XIII. Effects of Sulfates on Charge and Discharge of Sulfated Negative Plates

Sp. Gr.	Rate, Amperes	Charge	Discharge
Very low	low	U*	F† or E‡
Very low	high	U	F or U§
Low	low	U	F or E
Low	high	U	U
Moderate	low	U	U or E
Moderate	high	U	U

* U=unfavorable.

† F=favorable.

‡ E=equal.

§U=becoming unfavorable with cycling.

Viscosity and Electric Resistivity

Data on viscosity at 25 degrees centigrade are given in Table XIV; those for electric resistivity in Table XV. For viscosity the acid-salt solutions refer to data for 24.4±1.6 grams of the mixture per liter of acid; for electric resistivity to data for 23.0±0.7 grams of the mixture per liter of acid. A mixture of magnesium and sodium sulfates increases the viscosity of sulfuric acid solutions by about 6.6 per cent for specific gravities from 1.050 to 1.300 and by 8 per cent for the more dilute solutions. Since the rate of diffusion of ions is inversely proportional to the viscosity of the medium in which they exist, the observed increase in viscosity of sulfuric acid solutions resulting from additions of the sulfates of sodium and magnesium means that such additions decrease the rate of ionic movement whereby electrode polarization would be accentuated and electric watt efficiency lowered.

The mixture of sodium and magnesium sulfates increases the resistivity of sulfuric acid solutions by about 5 per cent for solutions exceeding about 1.2 per cent by weight of sulfuric acid. Only for extremely dilute solutions (less than 1 per cent, specific gravity below 1.01) is sulfuric acid containing a mixture of sodium and magnesium sulfates less resistant than pure aqueous sulfuric acid and this range of specific gravity is not encountered in the normal operation of a battery. This increase in the resistance of the electrolyte, although small, will

Table XIV. Data on Viscosity at 25 Degrees Centigrade of Sulfuric Acid Solutions and of Similar Solutions Containing Sodium and Magnesium Sulfates

Acid, Sp. Gr.	Grams M* per 1,000 Milliliters of Acid Solution	Absolute Viscosity, Centipoises	
		Pure Aqueous Sulfuric Acid	Pure Aqueous Sulfuric Acid + M*
1.009	21.8	0.921	0.996
1.050	22.5	1.04	1.11
1.102	25.8	1.22	1.30
1.197	26.7	1.66	1.77
1.300	25.3	2.41	2.56

* Mixture of sodium and magnesium sulfates.

Table XV. Data on Electric Resistivity at 25 Degrees Centigrade of Sulfuric Acid Solutions and of Similar Solutions Containing Sodium and Magnesium Sulfates

Acid, Sp. Gr.	Grams M* per 1,000 Milliliters of Acid Solution	Resistivity, Ohm-Centimeters	
		Pure Aqueous Sulfuric Acid	Pure Aqueous Sulfuric Acid + M*
1.0025	22.8	56.50	36.38
1.009	25.3	15.43	16.14
1.049	23.1	3.069	3.320
1.104	21.4	1.666	1.765
1.201	22.7	1.225	1.294
1.297	23.0	1.295	1.367

* Mixture of sodium and magnesium sulfates.

nevertheless affect to a small degree the discharges of batteries at high rates such as are encountered in the cranking of an automobile battery. Increased electric resistance of the electrolyte leads to increased internal resistance of the battery whereby it will operate at lower working voltages on high-rate discharges. This effect has been observed on a number of automobile batteries when subjected to 300-ampere discharges at 80 degrees Fahrenheit.

The increase in both viscosity and electric resistivity of aqueous solutions of sulfuric acid produced by the addition of a mixture of sodium and magnesium sulfates accounts, at least in part, for the differences in the effects noted previously for low and high charging and discharging rates for sulfated plates.

Effect of Magnesium and Sodium Sulfates on Discharge of New Batteries

In the foregoing sections we have discussed the effect during charge and discharge of a mixture of sodium and magnesium sulfates on hard sulfated negative plates which initially took the charge inefficiently. We shall now consider the effect of such mixtures on the discharge of new nonsulfated plates, or on plates that can be readily charged. In this, then, we are concerned primarily with comparing the efficacy of pure electrolyte to support the discharge with the efficacy of pure electrolyte containing a mixture of the salts.

DATA ON A $NA_2SO_4-MGSO_4$ BATTERY

To aid in interpretations of the phenomena that occur in studies of the efficacy of battery electrolyte to support discharges, tests were first made on a new and freshly charged automotive battery of 110-ampere-hour capacity in which the normal acid electrolyte was replaced by a nearly saturated solution of a mixture of sodium and magnesium sulfates. This battery was emptied of its normal acid electrolyte, was filled with distilled water, allowed to stand in this condition for an hour or two, and was then emptied. It was then filled with water and emptied 10 additional times over a period of 36 hours. Finally it was filled with a nearly saturated aqueous solution of a mixture of sodium and magnesium sulfates having a specific gravity of 1.220 and discharged at 5 amperes to zero voltage. This discharge is shown in Fig. 6. The discharge for a battery of the same type filled with 1.220 specific gravity sulfuric acid is also shown in the same figure for comparison. It will be seen that a battery can be made with a saturated solution of sodium and magnesium sulfates; however, this 3-cell battery operates at 1.5 volts over the main portion of its discharge period instead of at 5.9 volts as shown for a 3-cell battery filled with sulfuric acid solution of the same specific gravity. In other words, the working voltage of each cell of the Na_2SO_4–$MgSO_4$ battery is about 0.5 volt as compared to about 2.0 volts for a battery containing the usual sulfuric acid solution. The abrupt change in voltage that occurs in the discharge curve at 6 hours for the Na_2SO_4–$MgSO_4$ battery probably is a result of the residual acid retained by the plates.

It will be noted also that the Na_2SO_4–$MgSO_4$ battery gave slightly more hours to zero voltage than the H_2SO_4 battery but at a much lower voltage resulting in substantially lower total energy output. Here is, indeed, a most striking case where comparisons of battery outputs on an ampere-hour basis can yield erroneous conclusions. It is obvious by a glance at Fig. 6 that the Na_2SO_4–$MgSO_4$ battery gives only about 25 per cent of the watt-hour output or energy of the H_2SO_4 battery.

During these discharges the potentials of the plates were followed relative to a reference cadmium electrode. It was observed that a mixture of sodium and magnesium sulfates prolongs the discharge of the negative plate but at impracticable low cell voltages. It was found that the potentials of the positive plates were not

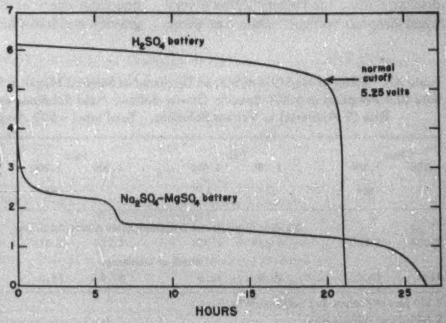


Fig. 6. Discharge voltages of a battery containing pure sulfuric acid solution (1.220 sp. gr.) and of a battery containing a sodium and magnesium sulfate solution (1.220 sp. gr.). Discharge rate 5 amperes

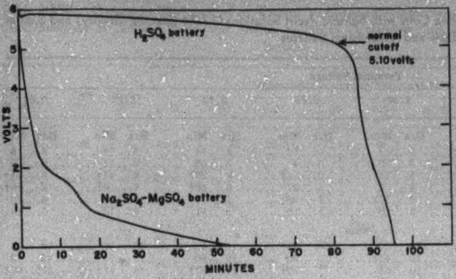


Fig. 7 (above). Discharge voltages of a battery containing pure sulfuric acid solution (1.220 sp. gr.) and of a battery containing a sodium and magnesium sulfate solution (1.220 sp. gr.). Discharge rate 50 amperes

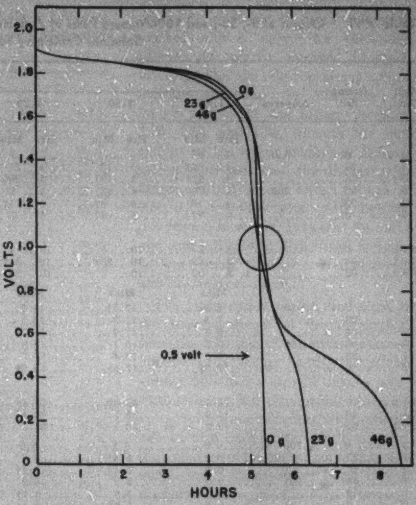


Fig. 8 (right). Effect of the addition of the indicated quantities of a mixture of sodium and magnesium sulfates on the discharge voltages of cells containing sulfuric acid solutions of 1.074 sp. gr. Discharge rate 5 amperes

improved but rather impaired owing to the insufficient supply of hydrogen ions needed for the discharge of positive plates (reverse of reaction 2). The low positive plate potentials observed during the discharge of the Na$_2$SO$_4$–MgSO$_4$ battery indicate, no doubt, that the electrode reaction which occurs at the positive plate of the Na$_2$SO$_4$–MgSO$_4$ cell differs from that which occurs at the positive plates in a cell containing pure aqueous sulfuric acid solutions.

In Fig. 7, comparable data are shown for a Na$_2$SO$_4$–MgSO$_4$ battery and a H$_2$SO$_4$ battery both of 1.220 specific gravity on 50-ampere discharges. It is evident that the former battery will not sustain high-rate discharges.

The discharges of acid-salt mixtures of the same specific gravity (1.220) on both the 5-ampere and the 50-ampere discharges will lie between the two extremes shown in Figs. 6 and 7.

Efficacy of Acid and Acid-Salt Electrolytes to Support Discharge

A new automobile battery was used in the study of the effect of the addition of a mixture of sodium and magnesium sulfates on the discharge of good nonsulfated plates. Each cell was provided with drain cocks as described in the section entitled "Changes in Specific Gravity During Charge," and after a full charge each cell was repeatedly filled with and drained of the acid electrolyte of required specific gravity to be studied. By this procedure it was assured that the sulfuric acid concentration of the solution in each cell was identical. Then cell 3 was again filled with the required volume of the sulfuric acid solution, cell 1 with the required volume of the same sulfuric acid solution in which 23 grams of the mixture of sodium and magnesium sulfates had been dissolved, and cell 2 with the required volume of the same sulfuric acid solution in which 46 grams of the mixture of sodium and magnesium sulfates had been dissolved. The cells were then ready for discharge.

The cell voltages and the plate potentials (measured relative to a reference cadmium electrode) were measured during discharge. Four series of discharges were made. The quantities of the mixture of Na$_2$SO$_4$–MgSO$_4$ added to the electrolyte of each cell in each series of discharges are given in the foregoing. In the first series the cells contained sulfuric acid solutions of 1.074 specific gravity and discharges were made at 5, 50, and 100 amperes; in the second series, sulfuric acid solutions of 1.190 specific gravity and likewise discharges were made at 5, 50, and 100 amperes; in the third series, solutions of specific gravities of 1.040, 1.074, 1.119, 1.160, and 1.190 were used and all discharges were made at 50 amperes. In the fourth series of discharges the electrolyte in each cell regardless of the quantity of the mixture that it contained was adjusted to a specific gravity of 1.220 and the cells were discharged at 50 amperes.

These discharges were then repeated in which 46 grams of the sulfate mixture were added to cell 3, and 23 grams to cell 1, and none to cell 2. Cell 1, therefore, served as control in both series of experiments. In Tables XVI through XIX outputs of each cell for four series of discharges are given to the indicated closed-circuit voltages. The sums of the outputs at each rate of discharge for all discharges are also shown for cells 3 and 2 with no salt, for cell 1 with 23 grams, and for cells 2 and 3 with 46 grams of salt. Therefore any inherent variations in the cells do not affect the summation of the outputs.

To facilitate the comparison of the outputs of the cells with or without the sulfate mixture the maximum time of dis-

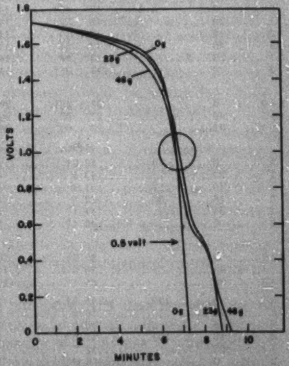


Fig. 9. Effect of the addition of the indicated quantities of a mixture of sodium and magnesium sulfates on the discharge voltages of cells containing sulfuric acid solutions of 1.074 sp. gr. Discharge rate 100 amperes

Table XVI. Output at 5-, 50-, and 100-Ampere Rates of Automotive Cells with Sulfuric Acid Solutions (1.074 Specific Gravity) and Similar Solutions Containing Sodium and Magnesium Sulfates

Cell No.	Grams M*	Amperes	Terminal Voltage 1.75	1.50	1.25	1.00	0.75	0.50	0.25	0.00
			Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min
3	0	5.0	4...21	5... 3	5... 7	5...10	5...12	5...14	5...16	5...18
1	23		4...22†	5...17	5...24	5...30	5...41	6...03	6...15	6...20
2	46		4...22†	5...19†	5...28†	5...35†	5...53†	7...27†	8...41†	9...04†
3	46	5.0	4...00	4...53	5... 2	5... 9	5...27	6...50†	8...6†	8...26†
1	23		4...10	5... 3†	5...11†	5...17†	5...28†	5...57	6...14	6...20
2	0		4...13†	5... 3†	5... 7	5...10	5...12	5...14	5...16	5...18
						Sum				
3+2	0		8...34†	10... 6	10...14	10...20	10...24	10...28	10...32	10...36
1+1	46		8...32	10...20†	10...35†	10...47†	10...69	12...00	12...20	12...40
2+3	92		8...22	10... 2	10...30	10...44	11...20†	14...17†	16...47†	17...30†
			Mins	Mins	Mins	Mins	Mins	Mins	Mins	Mins
3	0	50	8.2†	17.5†	18.2	18.4	18.7	19.1	19.5	19.8
1	23		7.3	17.3	18.6	19.6†	20.6	22.5	24.1	24.9
2	46		6.6	17.4	18.7†	19.6†	20.9†	24.2†	27.6†	29.3†
3	46	50	5.7	15.7	17.2	18.1	19.0	20.7	22.4	23.4
1	23		6.6	16.4	17.7	18.6†	19.5†	21.4†	23.0†	23.7†
2	0		7.2†	17.5†	18.2†	18.5	18.9	19.9	20.8	21.2
						Sum				
3+2	0		15.4†	35.0†	36.4†	36.9	37.6	39.0	40.3	41.0
1+1	46		13.9	33.7	36.4†	38.2†	40.1†	43.9	47.1	48.6
2+3	92		12.3	33.1	35.9	36.7	39.9	44.9†	50.0†	52.7†
3	0	100	0.0	5.2	5.6	6.0	6.1	6.2	6.4	6.5
1	23		0.0	5.4†	6.3	6.8	7.2	7.7	8.1	8.5
2	46		0.0	5.4†	6.4†	7.0†	7.4†	8.1†	9.4†	10.2†
3	46	100	0.0	5.2	6.2	6.7	7.2	7.7	8.6†	9.2†
1	23		0.0	5.5	6.4†	6.9†	7.3†	7.8†	8.4	8.8
2	0		0.0	5.8†	6.4†	6.6	6.8	6.9	7.1	7.3
						Sum				
3+2	0		0.0	11.0†	12.0	12.6	12.9	13.1	13.5	13.8
1+1	46		0.0	10.9	12.7†	13.7†	14.5	15.5	16.5	17.3
2+3	92		0.0	10.6	12.6	13.7†	14.6†	15.8†	18.0†	19.4†

* Mixture of sodium and magnesium sulfates.

† Indicates longest time of discharge to indicated voltage.

Table XVII. Output at 5-, 50-, and 100-Ampere Rates of Automotive Cells with Sulfuric Acid Solutions (1.190 Specific Gravity) and Similar Solutions Containing Sodium and Magnesium Sulfates

Cell No.	Grams M*	Amperes	Terminal Voltage 1.75	1.50	1.25	1.00	0.75	0.50	0.25	0.00
			Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min	Hrs Min
3	0	5.0	13...45†	14...43†	14...48†	14...52	14...57	15...02	15...08	15...12
1	23		13...37	14...38	14...48†	14...57†	15...08†	15...34	16...03	16...13
2	46		13...15	14...05	14...16	14...24	14...43	16...05†	17...12†	17...31†
3	46	5.0	12...54	13...57	14...06	14...14	14...28	15...00	15...51†	16...11†
1	23		13...26	14...25	14...34	14...42	14...53†	15...23†	15...48	15...57
2	0		13...35†	14...33†	14...38†	14...43†	14...51	15...09	15...15	15...19
						Sum				
3+2	0		27...20†	29...16†	29...26†	29...35	29...48	30...11	30...23	30...31
1+1	46		27...03	29...03	29...22	29...39†	30...01†	30...57	31...51	32...10
2+3	92		26...09	28...02	28...22	28...38	29...11	31...05†	33...03†	33...48†
			Mins	Mins	Mins	Mins	Mins	Mins	Mins	Mins
3	0	50	42.9†	52.7†	53.4†	53.8†	54.2	55.1	56.1	56.5
1	23		42.2	51.1	52.2	53.2	54.3†	56.8	59.2	60.2
2	46		41.6	50.4	51.2	52.5	54.1	58.8†	64.4†	66.2†
3	46	50	42.5	49.9	51.2	52.3	53.8	57.0	59.5	61.0
1	23		44.1	51.3	52.4	53.4	55.0†	59.7†	62.9†	63.7†
2	0		46.4†	52.8†	53.2†	53.6†	54.2	56.9	61.6	63.5
						Sum				
3+2	0		89.3†	105.5†	106.6†	107.4†	108.4	112.0	117.7	120.0
1+1	46		86.3	102.4	104.6	106.6	109.3†	116.5†	122.1	123.9
2+3	92		84.1	100.3	102.4	104.8	107.9	115.7	123.9†	127.2†
3	0	100	15.9†	22.2†	22.3†	22.5†	22.8	23.2	24.4	25.1
1	23		15.2	21.2	21.7	22.2	22.7	23.8	26.2	27.4
2	46		15.1	20.7	21.6	22.4	23.2†	25.2†	28.6†	30.4
3	46	100	12.0	18.4	19.2	19.8	20.4	21.7	23.5	24.4
1	23		12.7	18.7	19.3	19.7	20.6†	21.8†	24.7†	25.8†
2	0		13.9†	19.4†	19.6†	19.9†	20.2	20.9	23.0	24.4
						Sum				
3+2	0		29.8†	41.6†	41.9†	42.4†	43.0	44.1	47.4	49.5
1+1	46		27.9	39.9	41.0	41.9	43.3	45.6	50.9	53.2
2+3	92		27.1	39.1	40.8	42.2	43.6†	46.9†	52.1†	54.8†

* Mixture of sodium and magnesium sulfates.

† Indicates longest time of discharged to indicated voltage.

Table XVIII. Output at 50 Amperes to Various Voltages of Automotive Cells with Various Concentrations of Sulfuric Acid and Similar Solutions Containing Sodium and Magnesium Sulfates

Sp. Gr. H_2SO_4	Cell No.	Grams M*	Minutes to Indicated Terminal Voltage 1.75	1.50	1.25	1.00	0.75	0.50	0.25	0.00
1.190	3	0	42.9†	52.7†	53.3†	53.8†	54.2	55.1	56.0	56.5
1.190	1	23	42.2	51.2	52.2	53.2	54.3†	56.8	59.3	60.2
1.190	2	46	41.6	50.4	51.4	52.5	54.1	58.8†	64.4†	66.2†
1.190	3	46	42.5	49.8	51.2	52.4	53.9	57.0	59.5	61.0
1.190	1	23	43.9	51.3	52.4	53.3	55.0†	59.5†	62.8†	63.3†
1.190	2	0	46.3†	52.8†	53.3†	53.5†	54.2	57.0	61.5	63.5
					Sum					
	3+2	0	89.2†	105.5†	106.6†	107.3†	108.4	112.1	117.5	120.0
	1+1	46	86.1	102.5	104.6	106.5	109.3†	116.3†	122.1	124.0
	2+3	92	84.1	100.2	102.6	104.9	108.0	115.8	123.9†	127.2†
1.160	3	0	35.0†	43.1†	43.8†	44.2	44.7	45.7	47.1	47.8
1.160	1	23	34.3	42.4	43.5	44.4†	45.6†	48.5	51.6	52.5
1.160	2	46	33.2	41.5	42.7	43.8	45.3	50.6†	56.0†	57.8†
1.160	3	46	32.0	40.1	41.6	42.6	44.0	46.4	48.4	50.0
1.160	1	23	34.2	41.3	42.4	43.5	44.8†	47.9†	50.5†	51.4†
1.160	2	0	35.1†	43.0†	43.4†	43.8†	44.2	46.3	48.5	49.6
					Sum					
	3+2	0	70.1†	86.1†	87.2†	88.0†	88.9	92.0	95.6	97.4
	1+1	46	68.5	83.7	85.9	87.9	90.4†	96.4	102.1	103.9
	2+3	92	65.2	81.6	84.3	86.4	89.3	97.0†	104.4†	107.8†
1.119	3	0	24.0†	33.5†	34.0†	34.4†	34.9	35.9	37.2	37.6
1.119	1	23	23.0	32.1	33.2	34.3	35.5	38.0	40.0	41.0
1.119	2	46	22.1	31.3	32.8	34.0	35.8†	40.0†	43.8†	45.4†
1.119	3	46	20.0	28.5	30.0	31.0	32.2	34.9	37.0	38.4
1.119	1	23	21.4	29.4	30.5	31.5	32.7	35.8†	38.6†	39.7†
1.119	2	0	23.3†	31.7†	32.1†	32.5†	32.9†	34.7	36.9	37.6
					Sum					
	3+2	0	47.3†	65.2†	66.1†	66.9†	67.8	70.6	74.1	75.2
	1+1	46	44.4	61.5	63.7	65.8	68.2†	73.8	78.6	80.7
	2+3	92	42.1	59.8	62.8	65.0	68.0	74.9†	80.8†	83.8†
1.074	3	0	8.2†	17.5†	18.1	18.4	18.7	19.1	19.5	19.8
1.074	1	23	7.3	17.3	18.6†	19.5	20.6	22.5	24.1	24.9
1.074	2	46	6.6	17.3	18.6†	19.6†	20.9†	24.2†	27.6†	29.2†
1.074	3	46	5.7	15.7	17.2	18.1	19.0	20.7	22.4	23.3
1.074	1	23	6.6	16.4	17.7	18.6†	19.5†	21.4†	23.0†	23.7†
1.074	2	0	7.2†	17.5†	18.1†	18.5	18.9	19.6	20.7	21.2
					Sum					
	3+2	0	15.4†	35.0†	36.2	36.9	37.6	38.7	40.2	41.0
	1+1	46	13.9	33.7	36.3†	38.1†	40.1†	43.9	47.1	48.6
	2+3	92	12.3	33.0	35.8	37.7	39.9	44.9†	50.0†	52.5†
1.040	3	0	0.0	8.3†	8.6	8.8	9.0	9.2	9.4	9.6
1.040	1	23	0.0	7.3	8.7†	9.3	9.7	9.9	10.2	10.3
1.040	2	46	0.0	6.9	8.4	9.4†	10.3†	11.1†	11.6†	12.2†
1.040	3	46	0.0	7.2	8.5	9.4	10.4	10.6	15.0†	16.5†
1.040	1	23	0.0	7.6	8.9†	9.7†	10.5†	11.6†	12.6	13.1
1.040	2	0	0.0	8.0†	8.6	8.8	9.1	9.6	10.1	10.5
					Sum					
	3+2	0	0.0	16.3†	17.2	17.6	18.1	18.8	19.5	20.1
	1+1	46	0.0	14.9	17.6†	19.0†	20.2	21.5	22.8	23.4
	2+3	92	0.0	14.1	16.9	18.8	20.7†	21.7†	26.6†	28.7†

* Mixture of sodium and magnesium sulfates.
† Indicates longest time of discharge to indicated voltage.

Table XIX. Output (Minutes) to Various Voltages at 50-Ampere Rate of Automotive Cells with Electrolytes Adjusted to 1.220 Specific Gravity and Containing Sodium and Magnesium Sulfates

Sp. Gr. H_2SO_4 +M*	Grams M*	Cell No.	Sp. Gr. of Electrolyte	Terminal Voltage 1.75	1.50	1.25	1.00	0.75	0.50	0.25	0.00
1.220	0	3	1.220	51.0†	58.2†	58.6†	59.0†	59.6†	59.9	62.8	63.4
	23	1	1.220	47.7	54.9	55.8	56.8	58.0	61.0†	65.2	66.6
	46	2	1.220	45.1	52.3	53.5	54.7	56.4	60.8	67.4†	69.7†
	46	3	1.220	41.7	49.8	51.2	52.4	53.9	56.8	59.6	61.5
	23	1	1.220	47.8	54.7	55.9	57.0	58.4	62.3	66.3	67.6
1.220	0	2	1.220	52.4†	59.2†	59.5†	59.7†	60.3†	62.4†	67.6†	69.5†
						Sum					
	0	3+2	1.220	103.4†	117.4†	118.1†	118.7†	119.9†	122.3	130.4	132.9
	46	1+1	1.220	95.5	109.6	111.7	113.8	116.4	123.3†	131.5†	134.2†
	92	2+3	1.220	86.8	102.1	104.7	107.1	110.3	117.6	127.0	131.2
						Control (Pure Acid)					
1.220	0	3	1.220	49.6	56.2	56.6	57.1	57.5	59.7	62.6	63.3
1.220	0	1	1.220	50.4	56.3	56.7	57.1	57.6	60.4†	65.5	68.3
1.220	0	2	1.220	51.3†	56.9†	57.2†	57.5†	58.1†	60.3	67.3†	71.3†

* Mixture of sodium and magnesium sulfates.
† Indicates longest time to indicated voltage.

charge to each closed-circuit voltage is marked with a dagger for each discharge. It can be observed for each discharge that the cells containing no salt mixture sustained the discharge for longer periods at the higher voltages, 1.75, 1.50, and 1.25, whereas the cells containing the larger amount of the sulfate mixture sustained the discharge for longer periods at the lower voltage of 0.50, 0.25, and 0.00 volts. Although the differences observed between the acid and acid-salt cells at high cutoff voltages are very small, they are nevertheless real and can be reproduced.

The prolongation of the discharges at the lower voltages is inconsequential, since the cells free of the sulfate mixture operate at a higher voltage until the voltage falls to 1 volt, and thereby more watt minutes are obtained at operating voltages for those cells filled with aqueous sulfuric acid than for those cells containing a mixture of the sulfates of sodium and magnesium. This is shown in the typical discharge curves given in Figs. 8 through 11. Only at voltages below approximately 0.5 volt does a mixture of sodium and magnesium sulfates prolong the discharge. This value, 0.5 volt, is that discussed previously for the Na_2SO_4–$MgSO_4$ battery. Prolongation of the discharge at cell voltages below 0.5 volt by sodium and magnesium sulfates is of no practical value, and especially so since such mixtures lower the operating voltage in the practical voltage range above 1 volt. In each of these figures a circle is drawn around the discharge curves at 1 volt. It will be noted that the curves tend to cross or reverse their relative positions at about this voltage and any comparison of the relative merits of acid and acid-salt mixtures based on ampere-minute outputs to a terminal voltage of 1 volt obviously could be misleading. Comparisons should be based on watt-minutes or watt-hours, and not on ampere-minutes or ampere-hours. Preliminary discharges at 300 amperes of cells containing excess sulfuric acid, i.e., specific gravity of 1.280, have shown also that the addition of a mixture of sodium and magnesium sulfates to the electrolyte adversely affects the voltage characteristics in the practical range.

In Table XIX data are given on the output of cells all adjusted to a specific gravity of 1.220; one cell contained sulfuric acid solution only, another sulfuric acid solution to which 23 grams of the sulfate mixture had been added, and the third cell sulfuric acid solution to which 46 grams of the sulfate mixture had been added. The discharge results for these

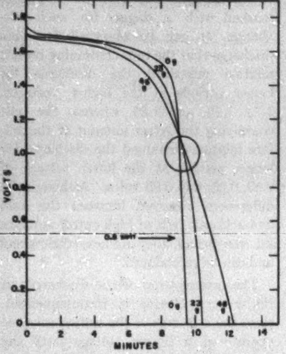
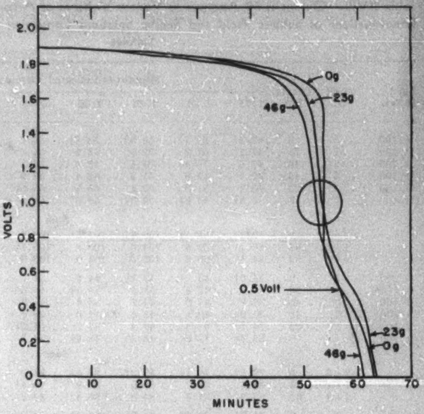


Fig. 10. Effect of the addition of the indicated quantities of a mixture of sodium and magnesium sulfates on the discharge voltages of cells containing sulfuric acid solutions of 1.040 sp. gr. Discharge rate 50 amperes

Fig. 11 (right). Effect of the addition of the indicated quantities of a mixture of sodium and magnesium sulfates on the discharge voltages of cells containing sulfuric acid solutions of 1.190 sp. gr. Discharge rate 50 amperes



cells containing solutions adjusted to a specific gravity of 1.220 show the same general phenomena as the results for solutions not adjusted to a common specific gravity; the superiority of the solution containing sulfuric acid only is, however, more marked.

Summary

The observations reported in this paper may be summarized as follows:

1. The specific gravities of the sulfuric acid electrolyte at the top and bottom of a lead-acid storage cell differ appreciably during charge and only attain a common value at the termination of charge when gassing ensues.

2. The rate at which hard lead sulfate on negative plates is converted to active material (lead and sulfate ion) during charging becomes less as the concentration of sulfuric acid is increased (or for higher specific gravities).

3. The conversion of hard lead sulfate on negative plates to active material (lead and sulfate ion) is inhibited during charging when mixtures of sodium and magnesium sulfates are added to the sulfuric acid electrolyte. This inhibition occurs for all concentrations of sulfuric acid solutions and for all charging rates. The detriment resulting from the addition of a mixture of sodium and magnesium sulfates to sulfuric acid solutions is more marked for dilute acid solutions and at higher charging rates.

4. The addition of mixtures of sodium and magnesium sulfates to sulfuric acid electrolytes of very low specific gravities has an apparent beneficial effect on the discharge of partially redeveloped sulfated plates. As the specific gravity increases on cycling the effect disappears. Mixtures of sodium and magnesium sulfates adversely affect the cycling of partially redeveloped sulfated negative plates at high rates for acid specific gravity over the range of 1.020 to 1.200.

5. Mixtures of sodium and magnesium sulfates do not significantly alter the solubility of lead sulfate in sulfuric acid solutions; in particular, measured differences are insignificant when compared to the solubility changes occurring in the solutions covering the normal operating range of a battery.

6. A mixture of sodium and magnesium sulfates increases the viscosity and electric resistivity of sulfuric acid solutions. The increase is 6 to 8 per cent for viscosity; for electric resistivity the increase is about 5 per cent.

7. A battery made with a saturated solution of a mixture of sodium and magnesium sulfates of specific gravity of 1.220 has an operating voltage of 1.5 volts on a 5-ampere discharge as compared with an operating voltage of 5.9 volts for a battery filled with sulfuric acid of the same specific gravity. The watt-hour output of a battery containing sodium and magnesium sulfates is only about 25 per cent of that obtained with a battery containing sulfuric acid solution.

8. A battery containing a saturated solution of a mixture of sodium and magnesium sulfates of specific gravity of 1.220 will not sustain a moderately heavy drain of 50 amperes.

9. Aqueous solutions of sulfuric acid of various specific gravities give more minutes of output to closed-circuit voltages of 1.75, 1.50, and 1.25 and operate at higher voltages than batteries filled with sulfuric acid solutions to which mixtures of sodium and magnesium sulfates have been added. The latter batteries give more minutes of output to closed-circuit voltages of 0.75, 0.50, and zero providing there is a deficiency of sulfuric acid. Although the differences in outputs observed at the higher voltages are very small they are nevertheless real and can be reproduced in new batteries.

References

1. On the Chemistry of the Plante and Faure Accumulators, J. H. Gladstone, A. Tribe. *Nature*, London, England, vol. 25, no. 221, 1882, p. 461. Also, *Electrician*, London, England, vol. 9, 1882, p. 612.

2. Chemical Reactions in the Lead Storage Battery, G. W. Vinal, D. N. Craig. *Report 778*, Journal of Research, National Bureau of Standards, Washington, D. C., vol. 14, 1935, p. 449.

3. Thermodynamic Properties of Sulfuric-Acid Solutions and Their Relation to the Electromotive Force and Heat of Reaction of the Lead Storage Battery, D. N. Craig, G. W. Vinal. *Report 1294*, Journal of Research, National Bureau of Standards, Washington, D. C., vol. 24, 1940, p. 475.

4. The Molal Electrode Potentials and the Reversible Electromotive Forces of the Lead Accumulator from 0° to 60° Centigrade, H. S. Harned, W. J. Hamer. *Journal*, American Chemical Society, Washington, D. C., vol. 57, 1935, p. 33.

5. Storage Batteries (book), M. Arendt. D. Van Nostrand Company, Inc., New York, N. Y., 1928, p. 60.

6. Storage Batteries (book), G. W. Vinal. John Wiley and Sons, New York, N. Y., 3rd ed., 1945, p. 216.

7. The Solubilities of Sparingly Soluble Salts Using Large Volumes of Solvents. I. The Solubility of Lead Sulfate, R. B. Purdum, H. A. Rutherford, Jr. *Journal*, American Chemical Society, Washington, D. C., vol. 55, 1933, p. 3221.

8. Solubility of Lead Sulfate in Solutions of Sulfuric Acid, Determined by Dithizone with a Photronic Cell, D. N. Craig, G. W. Vinal. *Report 1165*, Journal of Research, National Bureau of Standards, Washington, D. C., vol. 22, 1939, p. 55.

9. The Ionization Constant and Heat of Ionization of the Bisulfate Ion from Electromotive Force Measurements, W. J. Hamer. *Journal*, American Chemical Society, Washington, D. C., vol. 56, 1934, p. 860.

10. Ueber den Nachweis von Schwer Metallen mit Hilfe von "Dithizon". H. Fischer. *Zeitschrift fuer angewandte Chemie*, Leipzig, Germany, vol. 42, 1929, p. 1025.

◆

No Discussion

A Reappraisal of the Economics of Railway Electrification: How, When, and Where Can It Compete with the Diesel-Electric Locomotive?

H. F. BROWN
FELLOW AIEE

R. L. KIMBALL
ASSOCIATE MEMBER AIEE

Synopsis: The diesel-electric locomotive has been universally substituted for the steam locomotive on most of the railroads in this country. Railway electrification has not had further application during this period. Many believe this to be an indication that the economies of the diesel over the steam locomotive apply equally to electric operation as well. This opinion is debated. Power costs, investment costs, fixed charges, maintenance costs, and other operating costs of both diesel-electric and electric operation are compared. Stress is laid on the greater rise in maintenance costs, with age, of the diesel-electric than those of the electric locomotive. It is believed that when the difference in all these costs is fully determined, electrification will again be applied to certain parts of the American railroads having good load factors where electric power is available because of its greater economy. Commercial frequency applied to the contact wire at higher voltage, and the rectifier locomotive, offer means of standardizing future railway electrification. Examples are cited of studies recently made for electrification of this type in South America, and for a hypothetical installation in this country, to illustrate cost comparisons brought out in the paper.

THE rapid change from steam locomotives to diesel-electric motive power during the past 10 years on all but one or two of the railroads of the United States has been revolutionary. This change is shown graphically in Fig. 1. Today there are more than 22,000 diesel-electric locomotives in service, and the number of steam locomotives has been reduced to approximately 17,000.

Electric operation, introduced first in 1895, and applied more extensively between 1905 and 1925, had one more major installation in 1938. The total number of electric locomotives then in service was approximately 800. Since 1938 there has been no further important growth in railway electrification in this country, and the number of electric locomotives has remained more or less constant since that date.

This fact, coupled with the steady and almost universal application of the diesel-electric locomotive since 1940, has caused most laymen and many skilled railroad operators and technicians to believe that electric operation is outmoded and that the economies of diesel-electric operation over steam operation apply equally to electric operation as well.

This paper debates that thinking and points out under what circumstances and by what factors electric operation not only can, but probably will, grow again and compete economically with the diesel-electric for general railway operation.

Brief Review of Existing Railway Electrification

Space does not permit a detailed review of the existing railway electrification in the United States today. Nearly all of this electric operation was installed to solve special operating conditions and problems for which the steam locomotive could no longer serve economically, if at all. Much was installed for new railway operations such as new underground or enclosed passenger terminals, designed especially for electric operation and to accommodate heavy suburban traffic.

It is true that the diesel has superseded electric operation on one or two of the shorter railroad electrification installations, and it is possible that it may take the place of electric operation on other short tunnel installations which are a part of long through runs. On most of these electric operations, however, conditions are such that the internal combustion type of locomotive, even with its reduced amount of smoke and gas, must still be excluded as was the steam, even though both these types perform the major portion of long runs entering the electrified territory. It will be pointed out later that even if the diesel could be allowed in some of these electrified terminals and territories, it probably could not compete economically with an existing electrification. To this extent, then, electric operation is already competing, and will continue to compete, with the diesel-electric just as it has in the past with steam.

All these electrifications, because of their special conditions of application, have been very difficult and expensive to install and to maintain. To keep costs to a minimum in each case, a number of various "systems" have been devised, which has somewhat confused the issue in making electrification more acceptable to the railroads. It is probable that none of these systems varied more than 10 per cent in their over-all costs. All have proved their economy over steam.

Of all the various systems installed in this country one finally became outstanding in its application, namely the 11-kv 25-cycle single-phase system. This system is used today on half of the existing major railway electrifications, and accounts for more than 75 per cent of all the electrified track mileage. This same system, at a lower frequency, is widely used also in Europe.

By contrast with electrification, the railroads have found it comparatively simple to finance and apply by easy stages a more or less standardized type of diesel-electric power to replace worn-out and less economical steam power for general operations. Does it seem possible or probable that electric operation, with

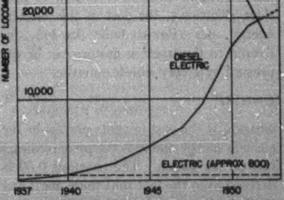


Fig. 1. The change from steam to diesel motive power in the United States

---

Paper 54-29, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 20, 1953; made available for printing December 7, 1953.

H. F. Brown and R. L. Kimball are with Gibbs and Hill, Inc., New York, N. Y.

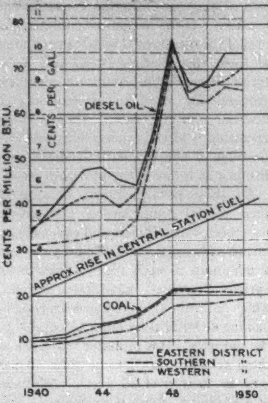


Fig. 2. Relative prices of railroad fuel reported to the Interstate Commerce Commission. Unit cost per Btu based on:

Coal: 12,000 Btu per pound
Diesel oil: 138,000 Btu per gallon

systems not yet standardized and with its handicap of a contact system, can compete with the diesel for this general railway operation? If so, how, when, and where can this be done? Obviously, all the factors must be carefully weighed.

Power Supply—Oil

The distribution of oil through pipe lines in the United States is as remarkable as the growth of the automotive industry and closely linked with it. No railroad has any difficulty in securing at any terminal adequate supplies of liquid fuel for diesel operation. The price of diesel oil, however, has increased during the past 10 years at a greater rate than the price of coal; see Fig. 2. Even so, the railroads are saving on their former fuel bill because of more efficient utilization of the fuel. Since all the diesel oil consumed for traction purposes on the diesel-electric locomotive is used to produce electricity, the cost of this electric energy, per kilowatt-hour (kw-hr), delivered to the traction motors can be expressed by a very simple equation

$$C=0.09P$$

where C is the cost in cents per kw-hr for electric energy and P is the price of diesel fuel oil per gallon. The constant 0.09 is derived from the fact that there are 3,413 Btu's in 1 kw-hr and 138,000 Btu's in 1 gallon of diesel oil, together with the assumptions that the diesel engine has an average thermal efficiency of 30.5 per cent and the electric generator with its auxiliaries has an average efficiency of 90 per cent. This constant favors the diesel somewhat as it is doubtful if this is always operating at this high efficiency at all loads and at all times between overhauls. This equation is shown graphically in Fig. 3 for diesel fuel at various prices.

It will be seen that electric energy generated on the diesel-electric locomotive takes about 12,400 Btu per kw-hr. This is comparable with the fuel rate of some not-too-modern central stations. With diesel oil at the current price of 10 cents per gallon, the cost of electric power on the locomotive is 9 mills per kw-hr. This is already higher than the "energy cost" of electric power for railroad traction power under some existing railroad contracts. This means that already, especially for "off-peak" operations, diesel-electric railway operation is costing more for fuel than is paid for electric power on some of the larger electrified railroads having general electric operation. This fact should be of interest to those railroad operators who would increase the utilization of idle diesel power by operating such into existing electrified territory in the place of electric locomotives. In most cases the cost of power alone for diesel operation would be greater than for electric operation.

Diesel oil will probably continue to be plentiful for some years to come. It is slowly rising in price, and there are no important factors in view to indicate a long-term change in this trend. It can be made readily available at any railroad terminal. It requires some transporting, storage, and handling facilities on the part of the railroads which add slightly to its cost on board the locomotive. These might be analagous to transmission and distribution losses for electric power.

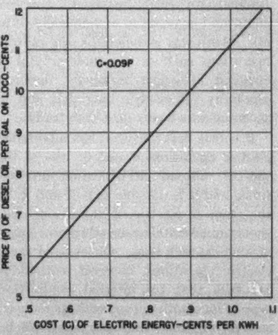


Fig. 3. Cost of electric energy produced on diesel-electric locomotive in cents per kw-hr

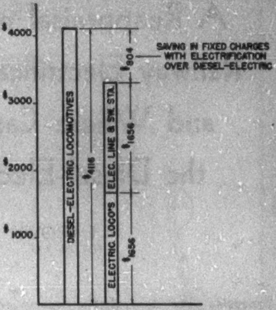


Fig. 4. Comparison of annual depreciation charges per $100,000 investment. Diesel-electric retired in 20 years, electrification retired in 40 years. Interest at 2 per cent, compounded annually

Power Supply—Electric

Many of the earlier railway electrifications were obliged to build and operate their own power plants because the commercial power stations were too small for their requirements. Even today some electrified railroads must continue to operate their own plants because there is no adequate supply of the required type of power available, and conversion equipment with purchased commercial power would be a greater expense.

Conditions, however, have changed materially during the past two decades relative to electric power generation. Large generating stations no longer are concentrated solely at or near large communities or industries. They are now being located nearer fuel sources, or at developed water-power sites, and are linked together by high-capacity transmission networks. Such networks, operating in synchronism, span not only entire states but groups of states in the more densely populated areas of this country. These networks and their supply stations have been increasing in area and in capacity at a very rapid rate during the past 10 years. During this growth the trend in the price of electric energy has been downward, in spite of its close relationship to the price of coal, which has had an upward trend; see Fig. 2.

Because of past limited supply capacity on the part of the power companies and the unattractive peak loads with poor load factors for the special railroad loads involved, power rates for railway electrification usually have included separate demand charges which appear high compared with the energy charge. On those railroads that have extended terminal electrification to include further operation of passenger, freight, and in some cases switching service, the load factor has been materially improved, sometimes above 60 per cent, and the demand charge, spread over more kw-hr, has been reduced to a smaller percentage of the energy charge and of the total unit cost per kw-hr. Such a load becomes more attractive to power supply companies, especially because of the diversity factor. General railway operations are as heavy at night as during the day, which gives a good diversity when combined with the usual industrial and commercial loads.

With the continuing growth and expansion of the electric power industry there will be an increasing probability that the general railroad loads, as distinct from the special railroad loads of the past and present, will become more and more attractive to the power industry as a valuable potential base load. This will be especially true if new sources for producing electric energy, possibly including atomic energy, should be developed. These, also, appear to be nearer realization than predicted 5 years ago. All these factors will tend to widen the present small difference in price per kw-hr at the locomotive between diesel-electric power and central station power for railroad operation. At present, spot checks indicate that for large blocks of power at load factors of above 60 per cent, there is already 1 mill difference between central station and diesel-electric power, in favor of central station power.

To avail itself of this power in the most efficient way, future railway electrification must use commercial frequency directly on the contact wire, and eliminate all expensive conversion equipment costs and losses. The means for doing this are now known and quite well developed.

First Costs and Fixed Charges

It has been comparatively easy to finance the change from steam to diesel power, a few units at a time, without drastic changes in the capital structure of the railroads. Electric locomotives can also be financed in the same manner, but of necessity the power supply and contact system must be financed and completed before any such locomotives can be operated. It is a sudden, wholesale change instead of a gradual substitution over a period of years. Electrification, in addition to the motive power, requires at present a different type of financing more difficult to negotiate than "equipment trust certificates." Nevertheless, if there should appear to be sufficient economic inducements in the future in favor of electrification, doubtless new financing methods could be devised for such new investments. The large financial institutions of this country are ever alert for promising investment opportunities. In some states, the taxation of railroad fixed property must be considered in the additional fixed charges.

All past railway electrification has been of a long-term investment by its nature. Even those installations which have been abandoned have had at least a 35-year life. This long life has applied to the locomotives as well as to the contact line and power supply. To place a life expectancy of at least 40 years for all the factors involved in electrification is justified by the large background of past experience.

The diesel-electric locomotive, on the other hand, cannot be expected to last this many years. By its very nature, containing a prime mover of the internal combustion type with all of its associated auxiliaries, and a multiplicity of precision-fitted reciprocating parts, constantly subjected to wear, it will of necessity have a much shorter life. Railroads today are allowed by the Interstate Commerce Commission to set up depreciation rates for diesel motive power on a 20-year life basis. There are indications that this should be reduced to 15 years in certain cases. This may be economical from the standpoint of more intensive utilization when compared with steam, but if compared with electric locomotives with the same high availability and utilization factors and a longer life, the fixed charges are going to be greater for the diesel by a ratio of two to one.

Diesel-electric locomotives currently cost between $100 and $110 per engine horsepower. This is equivalent to $125 to $135 per rail horsepower. There are indications that these prices may increase rather than decrease in the future due principally to the probability of future production in smaller quantities as the market becomes saturated.

Electric locomotives have been manufactured, currently, in much smaller quantities for about $125 to $135 per rail horsepower, equivalent to the diesel price. The trend in costs, however, for production if in larger quantities than at present, would be downward. Thus, motive power costs are more or less equivalent at present. Over the expected life of the electric locomotive, two diesel-electric locomotives must be purchased.

If the power supply equipment and contact system for a given railway electrification could be designed to cost no more than the cost of the motive power required, the total fixed charges over the life of the electrification would be less than the fixed charges for diesel-electric equipment for the same operation and

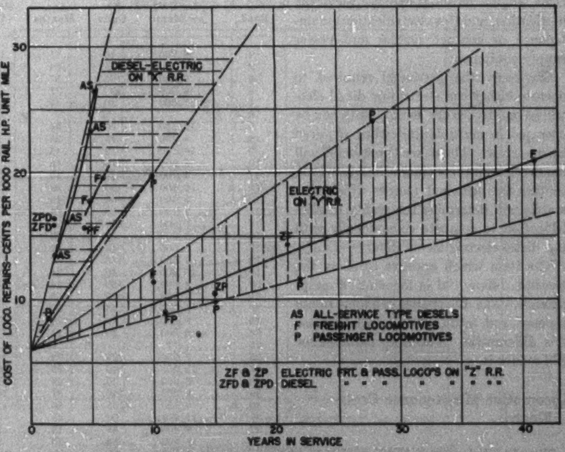


Fig. 5. Comparative cost of electric and diesel-electric locomotive repairs with age

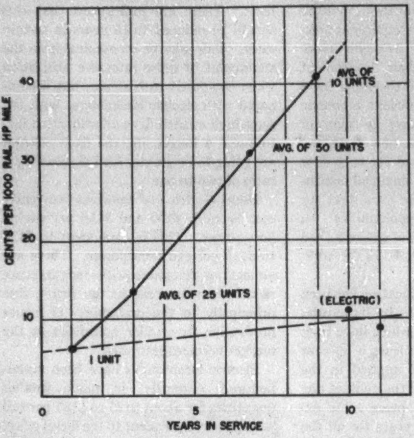


Fig. 6. Cost of diesel-electric road locomotive repairs classified by age. On one railroad. Units same size and manufacture. All costs made in 1950–51

period of time, as shown in Fig. 4. One criterion, therefore, to justify electrification is to insure that the contact system and power supply equipment do not exceed the cost of motive power for a given general railway operation. This would be one method of quickly determining whether or not an electrification project might be economical.

Operating Costs

Operating costs include, in addition to the power or fuel costs already discussed, wages of crew, locomotive maintenance costs, engine terminal expense including lubrication, and (for electrification) maintenance of contact system and power supply equipment.

Obviously the personnel required to operate either the electric or diesel-electric locomotive is the same. Both can be operated in multiple unit when required by one crew. The other three items will require careful consideration, for they contain the principal factors which ultimately will determine the ability of electrification to compete economically with the diesel-electric locomotive.

The item which appears to show the greatest differential is locomotive maintenance. The differential in engine house expense and lubrication would increase this differential. The line maintenance will reduce it.

Locomotive Maintenance Costs—Electric

The electric locomotive, being simply a power conversion machine and not a prime mover, has established the record over the past 45 years of operation of having the lowest maintenance cost of any form of motive power yet devised.

In order that the maintenance costs of locomotives of different weight and capacity and types may be equitably compared, it is necessary to have a common denominator of work performed. One way of doing this is to prorate maintenance costs over a unit-mile basis, the unit consisting of 1,000 horsepower measured at the rail. Thus, for a 4,000-horsepower electric locomotive the maintenance costs per locomotive-mile would be divided by 4 to give the costs for a 1,000-rail-horsepower unit-mile.

On this basis, current costs of maintenance of electric locomotives 40 years old have been investigated and found to be as low as 21 cents per 1,000-rail-horsepower unit-mile. Such costs, as well as current costs of maintenance of other electric locomotives of lesser age, on two important electrified railroads are shown in Fig. 5. It will be seen that these costs form a definite slowly rising pattern.

Locomotive Maintenance Costs—Diesel-Electric

Diesel-electric locomotives are still new. They have not yet had the years of operation that the steam or electric locomotives have had. Moreover, they are still being applied in such large numbers each year that it is very difficult to obtain, except in isolated cases, a true picture of the way maintenance costs will rise with the age of the individual unit. Many railroads keep records of maintenance costs simply on an entire class of equipment regardless of its age. Obviously with a continual influx of new units each year, the

Table I. Cost of Repairs of Diesel-Electric Locomotives per 1,000-Rail-Horsepower Unit-Mile

	Freight				Passenger			
Road	Avg. Miles per Unit per Month	Avg. No. Units	Avg. Age, Months	Cost in Repairs, Cents	Avg. Miles per Unit per Month	Avg. No. Units	Avg. Age, Months	Cost in Repairs, Cents
1	8,700	15	15	14.9				
2	10,500	25	64	11.5	18,500	15	31	7.35
3	9,550	105	57	10.1	16,600	20	67	12.1
4	9,000	20	37	8.35	11,000	25	85	14.2
5	11,000	20	42	11.3				
6					21,100	30	44	9.1
7	12,550	70	34	7.65	15,750	15	78	12.1
8	8,200	75	35	7.3	12,200	15	45	15.2
9	10,100	70	15	10.4	18,000	30	82	9.8
10	6,900	10	42	21.9				
11	4,700	55	54	27.0				
12	10,600	30	29	13.7				
13	9,900	60	40	12.0	13,750	20	19	9.6
14	7,100	55	24	10.95	7,100	25	23	10.95
15	7,500	270	25	14.45	19,300	45	43	15.6
16	5,900	55	46	21.1	10,300	20	20	15.15
17	8,000	15	22	10.9	11,000	20	21	8.6
18	9,050	130	32	13.7	16,800	30	30	7.23
19	12,300	40	34	7.37	12,000	25	24	10.32
20	8,000	450	24	10.05	22,400	50	48	11.4
21	9,200	620	53	11.7	23,900	145	57	7.75
22	7,550	245	30	12.2	17,200	65	91	14.3
23	7,200	80	74	23.3	11,600	30	58	14.5
24	8,000	30	21	11.2	12,800	15	46	9.2
25					7,900	10	53	13.66
26	7,000	200	22	10.9	15,800	50	64	13.4
27	12,200	190	59	7.1	16,100	90	65	8.4
28	6,900	30	49	17.7	9,600	10	49	15.3
29	9,000	150	37	10.54	17,500	75	35	14.2
30	9,550	155	36	11.6	16,400	65	69	13.5
31	11,600	60	29	9.4	20,750	30	28	7.9
32	8,400	85	15	8.95	24,250	50	52	8.3
33	8,800	85	20	8.6	13,700	25	26	11.5
34	8,100	90	59	10.9	11,650	40	42	9.2
Avg.	8,720		37	11.05	17,600		54	11.2

maintenance costs of such a class of equipment will be low and will remain low, or rise very slowly. It is only when a railroad has become entirely diesel-equipped, and has ceased to purchase new equipment in quantities, that it will begin to get a true picture of the rate of rise of maintenance costs.

Furthermore, there is at present a divergence of opinion as to just how such costs should be kept. One set of recently published maintenance costs covering several years shows these costs in "per cent of total operating costs." Obviously one has to know just how wages and fuel costs have changed over the same period before the true value of maintenance costs can be found.

Another proposal has been made that since the diesel engine will need periodic major repairs and replacements, the cumulative costs divided by the accumulated miles of service should be reported for each period. Obviously such a record would show eventually only the average cost of repairs over the life, and would not show the true rise, nor the point at which consideration should be given to replacement because of such high costs.

Occasionally it is possible to get maintenance figures for a group of locomotives classified by age. The data included herein are from authentic records of several railroads which were accessible. In Fig. 6 are shown costs of maintenance obtained in 1950–1951 of one type of diesel-electric locomotives used in combined passenger and freight service on one railroad, by age groups. All units were of same manufacture and capacity. All costs were within the same 12 months. The unit is reduced to 1,000-rail-horsepower unit-mile, the rail horsepower being taken as 80 per cent of the reported engine horsepower. The rise in costs is startling.

In Fig. 5 are shown, for comparison with the electric locomotive maintenance costs already referred to, current maintenance costs on another railroad of various types of diesel power, all reduced to the 1,000-rail-horsepower unit-mile basis. The comparison speaks for itself. Both railroads operate in nearly the same regional area and the service performed is comparable.

Because these are isolated cases, an analysis has been made of data, unpublished but authentic, compiled by one manufacturer from information furnished by more than 30 railroads, covering the costs of operation and maintenance of diesel-electric locomotives of different manufacture for the year 1951. These data are tabulated in Table I and are shown graphically in Fig. 7(A) for freight service and in Fig. 7(B) for passenger service. It will be seen that these data cover groups of locomotives of "average age" for each group. The results, as would be expected, are scattering, and lower than the costs shown in Figs. 5 and 6.

Nevertheless, the general trends indicate that diesel-electric maintenance, reduced to the same common denominator, will be for locomotives 10 years of age at least twice the maintenance costs of electric locomotives.

All of these last-mentioned data could be further subdivided into "engine repairs," "electrical repairs," and "other repairs," and the average percentage of total repair costs is as follows:

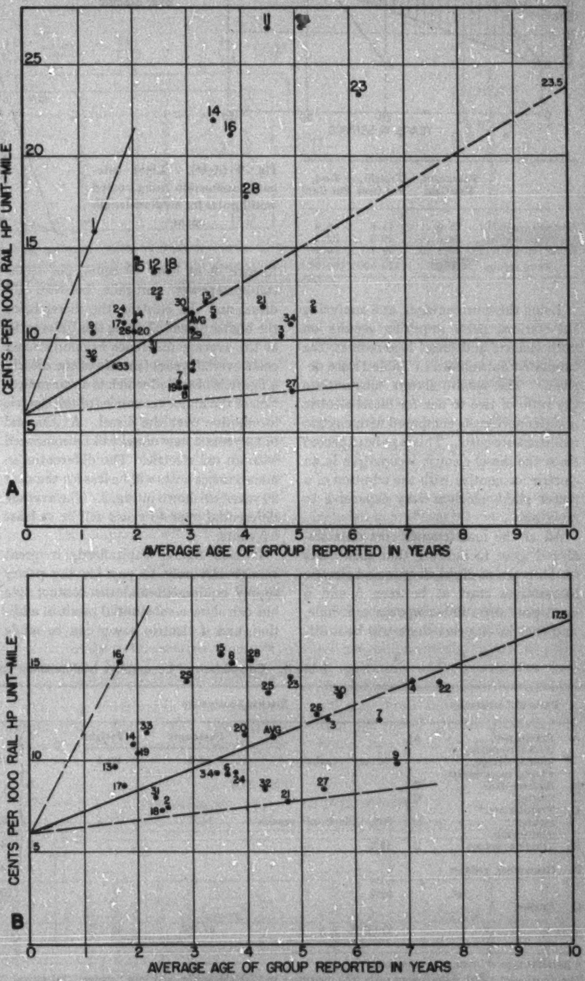


Fig. 7. Cost of repairs by groups of varying "average age"

A—For diesel-electric freight locomotives

B—For diesel-electric passenger locomotives

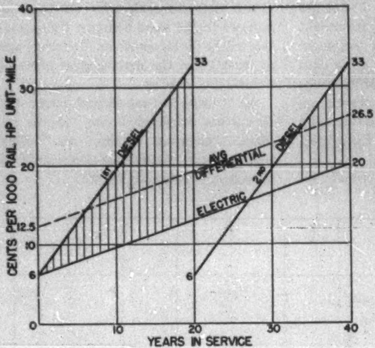


Fig. 8 (left). Difference in locomotive repair costs between electric and diesel-electric over life of electrification. Shaded areas indicate excess maintenance costs of diesel over electric

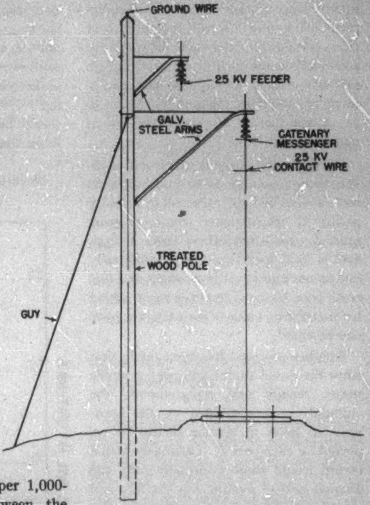


Fig. 9 (right). Light catenary construction using treated wood poles for medium-length spans

	Passenger, Per Cent	Freight, Per Cent	Yard, Per Cent
Engine repairs	33.0	41.6	48.9
Electrical repairs	24.5	29.3	22.2
Other repairs	42.5	29.1	29.2
Total repairs	100.00	100.00	100.00

Using these percentages, and analyzing the various parts requiring repairs on both electric and diesel locomotives, the tabulated data shown in Table II are derived. The results shown substantiate the ratio of two to one for diesel-electric maintenance costs compared with electric maintenance costs. This is to be expected since the diesel-electric locomotive is an electric locomotive with the addition of a power plant which is very expensive to maintain.

All of the maintenance cost data analyzed seem to indicate that such costs for all types of diesel-electric and electric locomotives start at between 5 and 6 cents per 1,000-rail-horsepower unit-mile, and that in 10 years there will be a difference of at least 10 cents, per 1,000-rail-horsepower unit-mile between the diesel and the electric, the diesel being the higher figure. Taking this differential as the average difference in maintenance costs over 20 years (the life of the diesel), a figure is obtained which is a fair indication of the major economies of the electric locomotive over the diesel. At the end of 20 years a new diesel will be compared with an old electric. The differential in maintenance costs will be less for the next 20 years, as shown in Fig. 8. The average differential over 40 years will be at least 6.5 cents.

If this figure, capitalized, is great enough not only to pay for the power supply equipment and the contact line but can show a substantial profit in addition, and if electric power can be made available, it is a logical conclusion that the electric locomotive must eventually replace the diesel. There are several minor items which can increase, and decrease, this major figure.

Engine Terminal Expense Including Lubrication

It is obvious that the diesel-electric locomotive needs more attention between runs than does the electric, cleaning, fueling, and lubricating being the principal difference. Of these items, lubricants alone are 0.7 cent per 1,000-rail-horsepower unit-mile, as given in the manufacturer's data referred to previously. The

Table II. Comparison of Locomotive Maintenance Costs

Parts of Locomotive	Electric Locomotive			Diesel-Electric Locomotive		
		Passenger	Freight		Passenger	Freight
A. Frame, cab, trucks, couplers, brakes, rigging, wheels, compressors	A } "Other" repairs	42.5	29.1	A } "Other" repairs	42.5	29.1
B. Blowers, fans	B			B		
C. Control	C			C		
D. Transformer,* rectifier, pantograph	D } 20% "Electrical" repairs	14.7	17.5	20% "Electrical" repairs	24.5	29.3
E. Traction motors	E 40%			E 40%		
F. Generators, exciters				F 40%		
	60%			100%		
G. Engine				G "Engine" repairs	33.0	41.6
	Total %	57.2%	46.6%		100.0%	100.0%
		Avg. 51.9%			100%	

* Rectifier type of locomotive assumed.

The division of diesel maintenance costs into passenger and freight service and into "engine", "electrical," and "other" costs are from same source as Table I.

Diesel "electrical" costs are divided into percentages shown and these percentages are applied to electric locomotive "electrical" costs, making these repair costs 60% of similar diesel locomotive costs.

The ratio of total diesel costs to electric costs is approximately two to one.

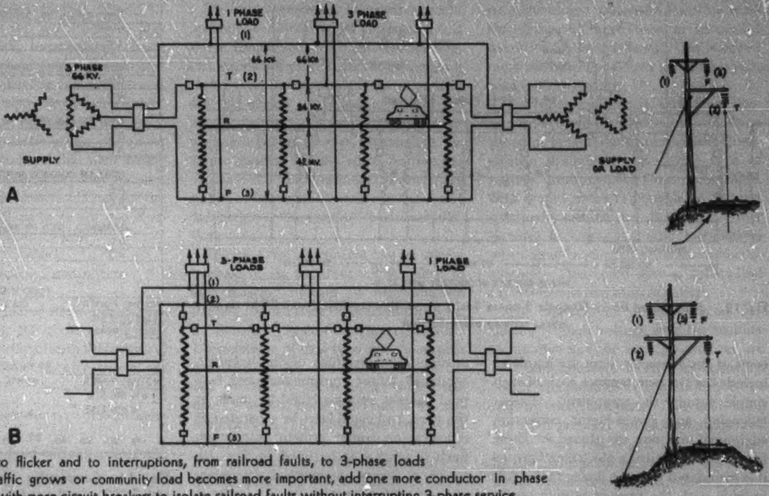


Fig. 10. Combined power and light traffic railway electrification scheme for developing new regions

A—Trolley and feeder are each one phase of 3-phase line. Phase 1 and feeder are supported on every other catenary supporting pole (indicated by the arrow at bottom right). Railroad right-of-way is power line right-of-way. Communities can be supplied with 1-phase or 3-phase power. Subject to flicker and to interruptions, from railroad faults, to 3-phase loads

B—As railroad traffic grows or community load becomes more important, add one more conductor in phase with trolley with more circuit breakers to isolate railroad faults without interrupting 3-phase service

other two items can easily increase this differential figure to 0.75 cent per 1,000-rail-horsepower unit-mile.

Future Contact Line Costs

The future contact line must be lighter and cheaper in cost than any previous construction, excepting possibly that used on the later a-c systems. This can be accomplished by the use of commercial frequency at a nominal voltage of 25 kv on the contact wire. Recent studies made for the possible electrification of two different railroads in South America using this system indicate that costs can be kept to within $18,000 per mile of track for such construction. Treated wood poles might be used in this country for much of such light construction; see Fig. 9. Such light construction can be designed to be as stormproof as any existing railroad contact line or transmission line.

Contact Line Maintenance Costs

The maintenance costs of the contact system always appears to be large on existing electrifications, since they are frequently for low load-factor operations. Such costs can be prorated either against the cost of power delivered to the locomotive (which is the most logical procedure), or against train mileage and added to the cost of locomotive maintenance, for comparison with other types of motive power.

On most of the existing electrified lines the current maintenance costs of the line, prorated as mentioned, are about

for d-c lines 2.5 to 2.6 mills per kw-hr
or 6.0 to 7.0 cents per train mile
for a-c lines 1.5 to 1.75 mills per kw-hr
or 4.0 to 5.5 cents per train mile

The wide variation in these figures reflects traffic density and load factor. They also reflect the very expensive type of contact systems installed. There is every reason to believe that lighter, less expensive types of contact systems will have lower maintenance costs, comparable to the ratio in first costs. This would give future maintenance costs of about

0.9 to 1.1 mills per kw-hr
or 2.5 to 3.0 cents per train mile

Obviously such costs would have to be included in any over-all comparison with diesel-electric operation.

Power Supply Equipment

With commercial frequency used on the contact system, the necessity for all conversion equipment disappears, as well as its operating and maintenance costs. Simple transformer and switching stations will step down and control the voltage from the power system network. The figures just given for the maintenance

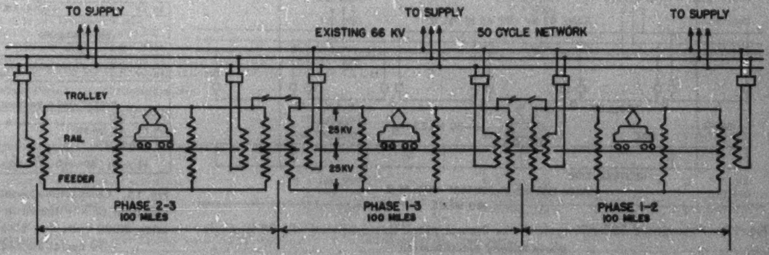


Fig. 11. A proposed 25-kv 3-wire single-phase 50-cycle railway electrification

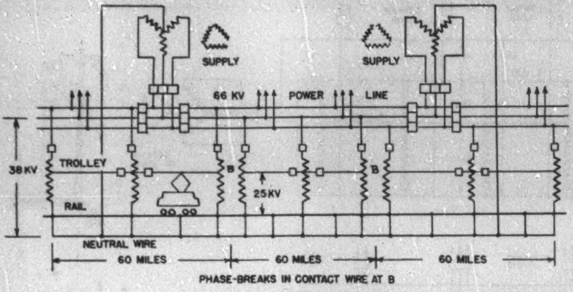


Fig. 12. A proposed 66-kv 50-cycle 3-phase 4-wire power line combined with 25-kv single-phase railway electrification

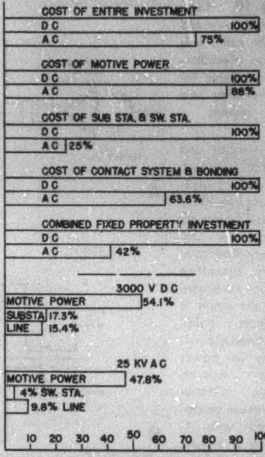


Fig. 14. Comparative electrification costs with 3,000 volts d-c and 25 kv, 50 cycles a-c for line shown in Fig. 11

costs of the contact system are ample to include the low maintenance costs of such simple supply tie equipment. Phase-balancing and power-factor correcting equipment are not envisioned in these future general railway electrifications, as it is considered that the power factor will be high with the type of motive power used, and that the load factor will also be high, and the load will be divided over all three phases of the power network. Phase-breaks in the contact system are nothing new, as they have been in use on the existing a-c lines for many years.

Phase-balancing, power-factor correction, inductive co-ordination, and elimination of harmonics cannot be ignored in special cases, nor should they be overstressed. All have been problems in the past, and not confined to any one system. The economical solution for all these problems is known, and if required must be included. They have not prevented the application of economical electrification in the past, nor will they in future.

Schemes to Make Electrification Attractive to Light Traffic Lines

In some cases, to develop new areas such as in South America or in Africa or possibly some sections of this country, combined power transmission and railroad contact systems could be made at the start, and could later be separated as industry or traffic growth warranted. Such a scheme is outlined in Fig. 10, where a 3-wire commercial frequency railway contact system and its feeder is part of a 3-phase line supplying power along the railroad. Such a scheme could readily utilize natural water power sources in countries lacking fuel, for railroad lines of even very light traffic, as the electrification would cost but little more than the power transmission system alone.

Other schemes suggest themselves. Fig. 11 shows a 3-wire scheme proposed for a South American railroad using power from an existing power network of more than 100,000-kw capacity. The railroad load would total about 30,000 kw and would be distributed over all three phases. Fig. 12 shows another scheme suggested for the operation of another South American railroad with lighter traffic. Here the transmission line would have to be constructed along with the railroad electrification, but would serve communities now without electric power of any kind. Power would be obtained from a relatively small thermal station. Although the railroad load would be distributed over all three phases, some phase-balancing equipment is indicated, as the railroad load is half of the supply capacity. Fig. 13 is another scheme studied for this same project using a higher trolley contact voltage, which eliminated all transformers.

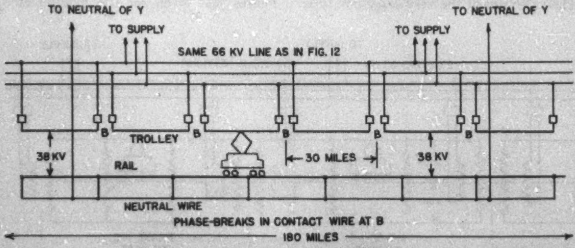


Fig. 13. A proposed 66-kv 50-cycle 4-wire 3-phase power line combined with 38-kv single-phase railway electrification

The comparative costs of the scheme shown in Fig. 11 with 3,000 volts d-c electrification are shown in Fig. 14. In Fig. 15 are shown the comparative costs

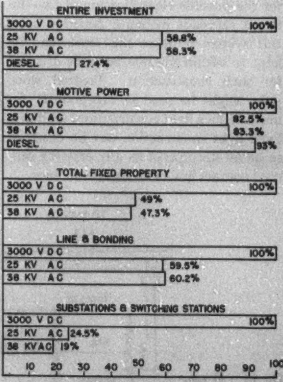


Fig. 15. Comparative investment costs. Electrification of line shown in Figs. 12 and 13, using 3,000 volts d-c, 25 kv a-c, and 38 kv, 50 cycles a-c and diesels

of the schemes shown in Figs. 12 and 13 with 3,000 volts d-c electrification and with diesel-electric locomotives. Both Figs. 14 and 15 clearly show the savings which may be realized in contact line and power supply equipment costs by use of the high-voltage commercial frequency system. In the latter study shown in Fig. 15, it was estimated that were diesels to be used the additional maintenance costs of the locomotives would pay for the additional capital costs of the a-c electrification in less than 13 years.

The Future Electric Locomotive

To enable the use of single-phase commercial frequency (60 cycles) on any future railway electrification in this country, two types of electric motive power are now available:

1. The motor-generator type of locomotive with d-c motors.
2. The rectifier type of locomotive, also with d-c motors.

Presumably the motors could be identical with the motors now used on diesel-electric equipment.

The motor-generator locomotive would probably be limited in its application to a single frequency. It might be possible, but somewhat more difficult, to design a motor-generator set to operate on both 25 and 60 cycles. A motor-generator set would certainly not be feasible for multiple-unit passenger cars, although in the future general railway electrification envisioned suburban passenger service may not be so important a factor in the traffic handled as in the existing special installations.

The rectifier locomotive has its appeal because of its apparent applicability to any existing system of electrification looking toward the eventual realization of a "Standard System of Electrification." Locomotives of this type are now under construction for one important electrified railroad in this country which will operate on 600 volts third rail in the terminal and on 11 kv 25 cycles a-c out on the main line.

It requires very little imagination to visualize some of the existing 25-cycle 11-kv systems extending for longer runs with 60 cycles 25 kv. The same transformers can be designed to operate on both these frequencies and voltages. Rectifier multiple-unit cars have been in successful operation for several years. One hundred more are nearing delivery for one a-c railroad in this country.

The future commercial frequency electric locomotive probably will be built to the same general designs as present diesel-electric equipment, with same type of motors, trucks, wheels, and underframe. This would make it possible to convert eventually some of the diesels into either rectifier or motor-generator locomotives when a major electrification is again considered. This could save a substantial investment in new motive power.

The application of commercial frequency at high voltage with the rectifier locomotive and the multiple-unit car are two comparatively recent developments which will do much to standardize future railway electrification. This will bring the fixed property investment costs well within the limits which can be financed by the savings to be realized in motive power maintenance, lower power costs, lower engine terminal expense, and lower fixed charges.

The railroads will profit from more

Table III. Electric Versus Diesel Operation of Railroad Represented in Figs. 16 and 17

Assumptions:
300 route miles of double track
650 miles of electrified track at $18,000 per mile
30 locomotives required, each 6,000-rail-horsepower
Cost of locomotives, $135 per rail-horsepower
Freight train miles per year, 3,720,000
Locomotive miles per year, 3,900,000
Depreciation rates:
  Electrification (40-year life) 1.656%
  Diesels (20-year life) 4.116%
Diesel oil: 10¢ per gallon (9 mills per kw-hr)

Max. power demand..................100,000 kw
Electric energy..........................531,000,000 kw-hr per year
Cost of electric power.................8 mills per kw-hr
Electric locomotive maintenance cost.....Avg. 10¢ per 1,000-rail-horsepower miles
....during first 20 years
Electric locomotive maintenance cost.....Avg. 13.5¢ per 1,000-rail-horsepower miles over 40-year life
Diesel locomotive maintenance cost......Avg. 20¢ per 1,000-rail-horsepower miles over 20-year life
Diesel lubrication oil differential........0.75¢ per 1,000-rail-horsepower miles

Immediate Investment	Electrification	Diesels	Difference in Favor of Electrification
Locomotives	$24,300,000	$24,300,000	
Contact line	11,700,000		
Power supply equipment	6,000,000		
Total investment	$42,000,000	$24,300,000	
Additional investment for electrification $17,700,000 (for first 20 years only)			
Average Annual Charges During First 20 Years			
Power	$ 4,250,000	$ 4,780,000	$ 530,000
Depreciation, locomotives	425,000	1,000,000	
Depreciation, fixed property	293,000		282,000
Locomotive maintenance	2,340,000	4,680,000	2,340,000
Diesel lubrication oil differential		175,000	
Line maintenance at 4¢ per locomotive mile	156,000		−247,000
Taxes and insurance on fixed property at 1.5%	266,000		
Total of all charges which will differ	$ 7,730,000	$10,635,000	$ 2,905,000
The average annual savings with electrification for the first 20 years will be $2,905,000			
This is an average gross return on the additional investment of $17,700,000 of 16.4% per year.			
Investment over 40-Year Period (Life of Electrification)			
Locomotives	$24,300,000	$48,600,000	
Fixed property	17,700,000		
Total investment	$42,000,000	$48,600,000	$ 6,600,000
Thus, for the long term, electrification will cost but 86.5% of the cost of diesels.			
Average Annual Charges over 40-Year Period			
All above charges other than locomotive maintenance	$ 5,390,000	$ 5,955,000	$ 565,000
Locomotive maintenance	3,160,000	4,680,000	1,520,000
Total of all charges which will differ	8,550,000	$10,635,000	$ 2,085,000

Electrification over its life will effect an ultimate saving in investment of $6,600,000, or 13.5% over the cost of diesels, and will effect average annual savings of $2,085,000 over diesel operation. This is an average gross annual return of 11.8% on initial additional investment.

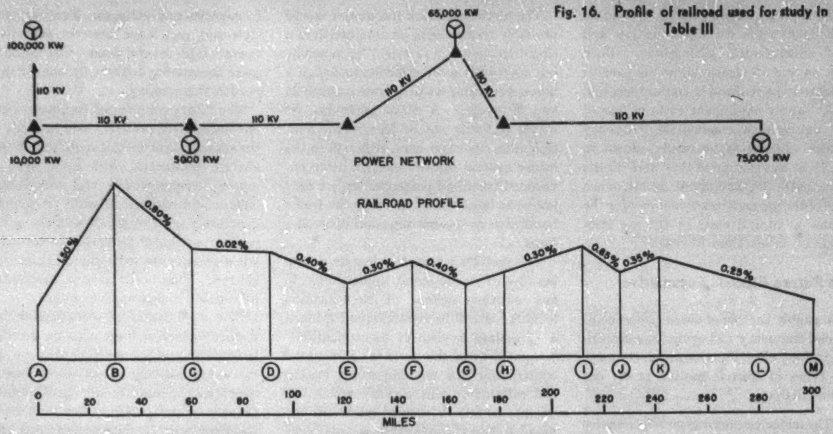

65,000 KW
100,000 KW
110 KV
110 KV
110 KV
110 KV
110 KV
110 KV
10,000 KW
5000 KW
POWER NETWORK
75,000 KW
RAILROAD PROFILE
1.50%
0.80%
0.02%
0.40%
0.30%
0.40%
0.30%
0.65%
0.35%
0.25%
A
B
C
D
E
F
G
H
I
J
K
L
M
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
MILES

Fig. 16. Profile of railroad used for study in Table III

economical operating costs. The electric equipment manufacturers will find new business to take the place of the falling off of the diesel-electric market, now under way. The electric power industry will profit from steady base loads having good load factors and diversity factors. The large investment institutions will find a good field for safe investment of large blocks of capital. The general public will gain improved transportation facilities.

Electrified railroads have always wanted to get out of the electric power generation business. All railroads are in this business now, in a more expensive way than they ever conceived they would be. Eventually all are going to look for ways and means of getting out of the diesel engine repair business, as well as the electric power generation associated with it.

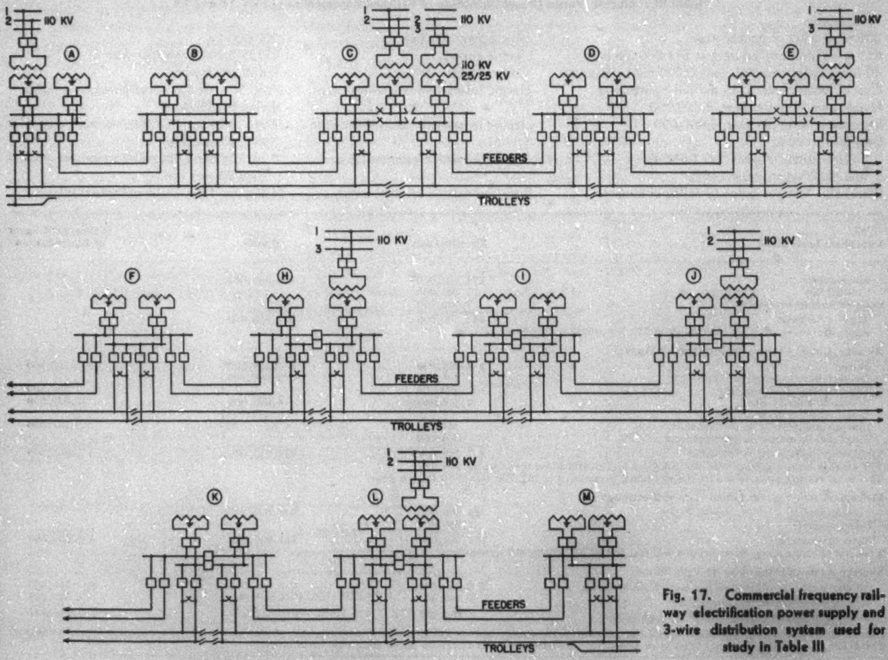

1
2
110 KV
A
B
C
D
E
1
2
3
110 KV
110 KV
25/25 KV
1
3
110 KV
FEEDERS
TROLLEYS
F
H
I
J
1
3
110 KV
1
2
110 KV
FEEDERS
TROLLEYS
K
L
M
1
2
110 KV
FEEDERS
TROLLEYS

Fig. 17. Commercial frequency railway electrification power supply and 3-wire distribution system used for study in Table III

Summary

How? Future railway electrification can compete with present diesel-electric operation chiefly through the ultimate savings to be realized in lower motive power maintenance costs, lower engine terminal costs, lower power costs, and lower fixed charges.

When? Serious attention will be given to a more economical type of motive power after the present diesel motive power has had at least 10 years of service and when the maintenance costs of the pool of this class of power are no longer kept at an artificially low level by the continued influx of a large number of new units each year. Also when the remaining steam locomotives are worn out and require replacement with new power.

A further impetus would be given by any application of studies now under way to utilize other energy sources to produce cheaper electric power.

Where? The first of these new installations will probably be made on the longer line railroads having a good traffic density and good load factor, in those districts served by adequate electric power networks and supplies. This does not necessarily mean water power, although such large developments could have an important influence on future railway electrification. It is also possible that some of the larger existing electrified lines, particularly those already operating on single-phase, may within the next decade decide to extend their electrification further. If so, commercial frequency and the rectifier locomotive will certainly be given careful consideration.

Conclusion

To illustrate all of these points, a study has been made of a hypothetical 300-mile section of a 2-track transcontinental railroad for electrification with 25 kv, 60 cycles single-phase, using the 3-wire system with 50 kv across the trolley-feeder. The section is assumed to include a heavy grade. Power is considered to be available from an adequate network, partly from water-power sources. The traffic consists principally of 17 freight trains per day in each direction, each requiring 6,000 rail-horsepower in motive power.

The profile of the line is shown in Fig. 16. Helper service is assumed on both sides of the heavy grade. The power supply equipment, transmission, and contact wires are shown in Fig. 17. For this study it is further assumed that the line is equipped with modern coded track signals. Other assumptions are detailed in Table III, which also shows the investment costs for electrification and diesel-electric locomotives, and those annual charges which would be different for the two types of operation compared.

This comparative study indicates that for the long term covered by the life of the electrification, the total investment required for electrification is actually 13.5 per cent less than that required for diesel-electric locomotives. The savings amount to $6,600,000.

For electrification an investment in addition to the cost of diesel locomotives is required for the first 20 years (the life of the first set of diesels) amounting to $17,700,000. The average savings in annual charges during this period will be $2,905,000, or an average annual gross return of 16.4 per cent on this additional investment. The standardization of electric locomotive manufacture could reduce the cost of electric locomotives in the order of 25 per cent over present costs. Applied to this study this would reduce the "additional investment" for electrification to $11,700,000, and the gross annual return would then be approximately 25 per cent.

◆

Discussion

R. L. Webb (Consolidated Edison Company, New York, N. Y.): The authors are complimented on having brought together in one paper a very comprehensive review of the diesel-electric and the electric locomotive engineering and economic comparisons. This discussion will deal with the "Power Supply Equipment" and "Schemes to Make Electrification Attractive to Light Traffic Lines" sections of the paper.

The electric supply company is, of course, interested in the effects the railroad load may have on this system from a phase balance, voltage stress or balance, and how these and the anticipated railroad load may affect the over-all capacity of its transmission lines for carrying 3-phase loads along the railroad right-of-way.

In Fig. 10 the 3-phase system is shown to be grounded through the rails at a point one-third of the way between the phase terminals of the 66-kv delta-connected winding. The system will certainly work, but it has the unfavorable feature of imposing unsymmetrical voltages on each of the phases to ground. This, of course, is undesirable from an insulation and lightning-arrester application viewpoint. If the line is long, or if it is made up of cable, the unbalanced charging currents, which will also result, may cause voltage unbalances that might cause further objection.

In all of the schemes covered in Figs. 10, 11, 12, and 13 there is the possibility that the connected rotating machinery including the generators may be required to carry a certain amount of negative-sequence current. An effort is made to avoid this in Figs. 11, 12, and 13, but it must be recognized that the railroad load in each of the track sections may vary and, if so, would impose some unbalance on the 3-phase system. A moderate amount of this can be carried without difficulty on most large systems, but this is a point which must be looked into carefully.

It is assumed that consideration has been given to a scheme of power supply to single-phase railroad electrification whereby a 3-phase to 2-phase transformation is utilized with capacitors and reactors, to take balanced 3-phase power from the power system while delivering single-phase power at the required voltage to the trolley system. This arrangement would permit tapping the 3-phase power system at any point and supplying the trolley at a voltage which is always approximately in-phase at all points on the railway electrification. Phase-balancing equipment, whether static or rotating, causes losses which must be evaluated in any scheme of electric supply.

The Consolidated Edison Company has supplied single-phase power to one of the railroad systems northeast of New York City for many years. The supply system is somewhat similar to a 1-phase section of Fig. 11. It has operated entirely satisfactorily to both parties except for some difficulties experienced by the power company with the amortisseur windings in the supply generators. This item of difficulty has now been overcome.

W. S. H. Hamilton (retired) (New York Central Railroad, New York, N. Y.): I agree most heartily with the authors that the proper system for general electrification of the future is one with 20 to 25 kv on the trolley at commercial frequency, i.e., 50 to 60 cycles. That, combined with ignitron rectifiers on the rolling stock, so that d-c traction motors are used, will solve practically all of the existing engineering problems. Such a system does not prohibit the use of motor-generator locomotives or locomotives having single-phase motors capable of operating on 50 to 60 cycles, wherever they may fit economically, or meet some particular requirement.

However, I do wish to disagree with some of the statements made by the authors, particularly in regard to the economics.

I disagree with the statement "Over the expected life of an electric locomotive, two diesel-electric locomotives must be purchased," found in the section on "First Costs and Fixed Charges." There is no reason why the running gear (mechanical parts) and traction motors of a diesel-electric locomotive should not last as long as the corresponding parts of an electric locomotive.

The life of the diesel engine itself is, of course, still open to question, but 15 years is a fairly accurate estimate of it. If the authors had said that over the life estimated for an electric locomotive two diesel engines would be required, I would agree. The cost of a diesel engine and accessories (and generator, which might or might not be salvaged) is approximately one-third the cost of a complete diesel-electric locomotive. On this basis the total investment for diesel locomotives as compared with electric locomotives would be $1^{1}/_{3}$ instead of two, as the authors state.

While the following information may not be entirely pertinent to the subject, attention is called to an article of mine[1] which describes the performance of the New York City 3-power locomotives after 22 years of operation. After 22 years' service, about 10 per cent of the cylinders are still as they came originally, about 75 per cent of the original cast iron pistons are still in service, and about 25 per cent of the original crankshafts are still in service. Crankshaft life (which is the largest single item in deisel-engine maintenance) has averaged approximately 12.5 years.

While these diesel engines are low horsepower (hp), relatively slow speed engines, they run at a fairly high load factor somewhat comparable with the diesel road locomotives of today. These locomotives were the largest single group of diesel-electric locomotives to be placed in service at the time they were built (1930–31).

It would seem from the performance of these and other diesel locomotives that a life of 15 years for the diesel engine, generator, and accessories could be safely assumed for an economic statement. What effect this will have on the conclusions arrived at by the authors I have not tried to estimate.

In the last paragraph of the section on "First Costs and Fixed Charges," I wish to point out that electrification usually means a change in the signal system, which is quite expensive and which ordinarily gets charged to the cost of electrification. This should be kept in mind in connection with the cost of the overhead construction.

In the following section I agree most emphatically with the authors that costs of maintenance should be kept on a 1,000-rail-hp unit-mile basis or some comparable unit. There have been many attempts to keep them on all sorts of bases recently, which add to the confusion of attempting to analyze and compare the results of operation. If the rail-hp unit-mile basis is used, it should be on the basis of the actual continuous rating that the traction motors (or generator, if it is limiting) are capable of delivering in accordance with AIEE Standards.

Referring to this same section, there should be a difference in cost of maintenance in favor of electric locomotives. It must be borne in mind that from this difference must be deducted the cost of maintenance of the overhead contact system, substations, etc., together with the increase (if any) in the cost of maintenance of the signal system. One must not overlook this in comparing the relative maintenance costs of electric and diesel-electric locomotives.

Referring to the following sections, I would like to ask if the maintenance cost for the overhead system on a per-train-mile basis are comparable with the costs of maintenance of the electric locomotives on a per-mile basis, so that they may be added to them for a quick comparison with the cost per mile of maintaining diesel-electric locomotives.

In the section on "The Future Electric Locomotive," the authors confute themselves to some extent as to the life of diesel-electric locomotives when they state, "This would make it possible to convert eventually some of the diesels into either rectifier or motor-generator locomotives when a major electrification is again considered." If the diesels have to be replaced every 20 years as stated previously, there would be little or no incentive for conversion. It would be cheaper to continue to operate them as diesels until their life was run out. On the other hand, if the life of the running gear and traction motors is considered as approximately the same as that of an electric locomotive, then there would be a strong incentive for conversion.

I wish to criticise Fig. 9, which shows a 25-kv trolley and a 25-kv feeder. The best information I have is that the feeder should be approximately twice the trolley voltage, or in this case 50 kv, in order to secure the maximum efficiency and the greatest distance between feed-in points. If this is done, the arm supporting the feeder should be on the side of the pole away from the catenary.

It is my understanding that with such a system, if the trolley voltage is taken at not more than 24 kv and the feeder voltage at 48 kv, which results in a 72-kv transformer with a ground 24 kv from one end, the change in insulation standards which occurs at 49 kv is avoided. Thus the cost for wayside or feeder transformers (as distinguished from those used at power feed-in points from the power system) is kept to a reasonable minimum and the distance between power feed-in points greatly increased, due to what amounts to a 72-kv transmission system.

Furthermore, in connection with Figs. 10–13 and Fig. 17, it is my opinion that in discussing general electrification these various arrangements should not be brought in, and that at the power feed-in points static conversion from 3-phase to single-phase, along the lines of that worked out by Battelle Memorial Institute, with which the authors are doubtless familiar, should be considered as the standard basic arrangement. The sooner everyone goes to work on this problem the quicker it will be solved.

It is undoubtedly true that some of the figures mentioned might be used in special cases, but it would take a peculiar set of circumstances for them. For instance, Fig. 11 shows loading on alternate phases. The difficulty with this arrangement is that the load can be 25 per cent on phase 2-3, 75 per cent on phase 1-3, and 0 per cent on phase 1-2 at any given time. This would require the power company to absorb considerable phase unbalance, which is not likely to be the case unless the railway load is quite low, and we all know from sad experience that lightly loaded railroads are not likely to electrify.

With the static 3-phase to single-phase transformation at all power feed-in points, this situation is avoided; and with a 24-kv trolley and a 48-kv feeder very few feed-in points will be required.

REFERENCE

1. N.Y.C. THREE-POWER LOCOMOTIVES, W. S. H. Hamilton. *Diesel Railway Traction*, London, England, vol. 7, no. 251, April 1953, pp. 77–80.

Dr. G. Huldschiner (College of the City of New York, New York, N. Y.): This paper is exceedingly interesting and challenging and there is no doubt that the mercury-arc rectifier locomotive will play an important role in future electrifications. There exists, however, another promising possibility to use a-c electrification with industrial frequency, by equipping the locomotives with industrial-frequency a-c commutator motors.

Such locomotives were built and successfully run on a small mountain railway line in Southwestern Germany, the Höllental-Bahn (Black Forest), presenting grades up to 4.4 per cent, trolley voltage 20,000 volts, and frequency 50 cycles (European industrial standard). These locomotives are of the B_0-B_0 type, output about 3,600 hp, hourly rating. The line is situated in the French occupation zone, and the French engineers became so much interested in this solution that they ordered experimental locomotives of an equivalent type and intend now to introduce this system for their secondary railway lines.

As compared with the rectifier locomotives, the use of industrial frequency a-c commutator motors would represent a substantial saving, and consequently shift the economical picture still more in favor of full electrification, through the elimination of first cost and maintenance for rectifiers and filters. The total weight on drivers, dictated by adhesion, would of course be the same, but the smaller construction weight of the a-c commutator motor locomotive could be brought up to the required figure by using cheap ballast. Moreover, the elimination of rectifier and filters would result in a shorter and less crowded locomotive-cab, an advantage which should not be underestimated.

Of course, the use of the American standard of 60 cycles would increase the difficulties of commutation, but as it was successfully possible to jump from $16^{2}/_{3}$ cycles (European standard for single-phase commutator railway motors) to 50 cycles, the step from 50 to 60 cycles should not be too onerous.

There are other problems connected with the use of industrial-frequency a-c commutator motors, but still it presents a very interesting and promising solution for full railway electrification, much more challenging than the equipment of locomotives with motor generators, a variety which had no past and will probably not have any future.

T. F. Perkinson (General Electric Company, Schenectady, N. Y.): There are two items of operating costs in the paper which I should like to discuss, since these two particular items exert a dominant influence in the over-all economic comparison between the two systems.

The first item is locomotive maintenance. The authors use 10 cents (¢) per 1,000-rail-hp per mile for electric locomotive maintenance. I presume that the type of locomotive considered is the rectifier type. Since this type of motive power has not been operated in other than experimental service in the U. S., we do not have actual operating costs on which to base comparisons. There are, however, actual operating figures available on the cost of maintenance of motor-generator locomotives in the same type of service considered in the paper.

Available statistics on motor-generator locomotives approximately 6 years old and making 100,000 miles per year show a stabilized maintenance cost of 4.26¢ per 1,000-rail-hp per mile. This cost does not include any heavy overhaul expense and it might be expected reasonably that the average maintenance cost would rise to something of the order of 5.5¢ per 1,000-rail-hp per mile for the 6,000-hp locomotive considered in the paper. On this basis the total annual motive-power maintenance cost for electric locomotives in the example considered in the paper would amount to $1,285,000 instead of $2,340,000 shown.

If the 2-to-1 ratio between diesel-electric and electric locomotive maintenance costs be allowed, the differential increase is $1,285,000 instead of $2,340,000. The total diesel-electric maintenance would then appear to be $2,570,000 instead of $4,680,000.

The figure of $2,570,000 divided by 3,900,000 locomotive miles results in 66¢ per locomotive mile, which amounts to 11¢ per 1,000 rail-hp per mile on the basis of a 6,000 rail-hp locomotive—which is the equivalent of a 7,200-hp locomotive—as diesel-electrics are commonly rated.

Diesel-electrics of smaller unit hp in the type of service under discussion are operating for costs approximately 11¢ to 13¢ per 1,000 rail-hp per mile and these are not exclusively new locomotives. With larger hp units the cost per 1,000-rail-hp mile should be somewhat lower than the 11¢ figure.

From the foregoing it would appear that the unit maintenance costs employed in the paper for both types of motive power are too high for the particular service considered in the hypothetical illustrative example.

The second item subject to discussion is the matter of power-fuel costs. Presumably the authors have assumed that electric energy will be secured for a cost of 8 mills per kw-hr delivered to the motors on the electric locomotive. With a transmission-distribution-conversion efficiency of 81 per cent, this indicates that the purchased price of electric power is expected to be approximately 6.5 mills per kw-hr as an average.

Obviously the availability and cost of electric energy are important considerations in any electrification study, and unless these are pretty well determined for a particular case, economic comparisons remain somewhat academical in nature.

If the contemplated electrification were to have available to it hydro-generated energy, the cost employed in the paper may be approximately correct, particularly in the Northwestern part of the United States. If steam coal-based generation is involved, the figure is too low in today's power market, particularly where new power-supply contracts are to be negotiated.

If the 8-mill energy cost used in the paper is intended to apply at point of purchase, the annual charge for power for the example selected will be approximately equal to the combined annual fuel-lubrication costs for equivalent diesel-electric service with fuel oil carrying a unit cost of 10¢ per gallon.

The matter of proper depreciation rates for use with the two types of motive-power is one that is open to debate. The diesel-electric is too new to permit of experienced age statements, whereas there are electric locomotives still in service in the U.S. with ages above 40 years. Some of these electric locomotives are still in use largely because economically superior types of electric locomotives applicable to the particular services and operating conditions involved have not been produced during this life span. Other electric locomotives have gone to the scrap pile after 25 to 30 years of life because they have been superseded by more economical types of electric or other types of motive power.

Current designs, and even designs produced 10 to 15 years ago, of diesel-electrics are still in service and will be in service many years hence, unless economically superior designs are brought out to replace them. In short, these locomotives will not wear out, because in the normal maintenance operations they are being constantly renewed and, in some cases, improved with new or repaired components.

Current designs will be replaced only when and if designs which will yield a substantial return on the required new investment become available.

The current clamor for permissible depreciation rates on diesel-electrics based on 15 to 20 years life is borne of the obvious advantages that accrue accounting-wise to the user where accelerated amortization rates, not based on actual life expectancy, are permitted. Accelerated amortization rates may wear out motive power on paper, but not necessarily in fact. There are electric locomotives still in service that have been worn out "on the books" but they still continue to operate in fact.

Unless new designs are evolved which will show appreciably lower operating costs, the railroads will continue to renew the older motive power—be it electric or diesel-electric—and to extend the life well beyond the 20 years assigned to the diesel-electric in the paper. If new designs are evolved, credit should be allowed in comparisons for the lower operating cost procurable with the replacing design that causes the shortened life of the superseded equipment.

Diesel-electrics 12 to 14 years of age are being overhauled, rewired or rebuilt today with the expectancy of retaining them in service for at least an equal number of years in the future—and possibly more.

As a general observation on the subject of railroad motive-power economics, it may be said that if electric energy is available at a low enough rate, if the traffic density is high enough, etc., electrification can be made to show an economic advantage over diesel-electric operation. However, the economic superiority of the former over the latter must be appreciably greater than a few percentage points before any railroad will consider justifiable the financial risk involved in the changeover. Too many of recent studies, where all of the influencing factors have been taken into account, show an advantage for electrification that is marginal at best, and this marginal superiority is not of sufficient magnitude to warrant the assumption of risk that accompanies the additional capital lock-up involved with electrification projects.

There is no disputing that the electric locomotive is the nearest to the ideal in motive power from a performance standpoint. Given a power-supply system the electric locomotive will outperform any other known type of motive power.

From a fuel-utilization viewpoint, the electric locomotive can, through stationary power-plant generation, use any of the prime energy sources currently known to industry and which might be developed in the future. Included in this category are such prime sources as water power, gas fuels, direct solar energy, and atomic power, which are not available for use with other than electric motive power.

The principal deterrent to a more widespread use of electric motive power in the U.S. is the relatively high capital investment required by the power-supply system—generation, conversion, and distribution facilities—and the difficulties (financial and operational) attendant to the installation of the power-supply system.

Railroad electrification stands, however, as an ever-available ultimate way of meeting the country's railroad motive-power requirements, and as a form of insurance that warrants the interim use of more-or-less economical—but less satisfactory, performance-wise—forms of motive-power that permit the direct on-board use of currently economically available fuels.

Evidence that electric motive power will eventually come into more general use in the U.S. is at hand in the circumstances existing in those countries where fuel oils are economically less available than in this country. Confronted with the hard facts that the use of fuel oils (and in some cases, coal) for their railroad motive power is currently restricted, and that the future holds negative encouragement that the situation will improve, several countries external to the U.S. have embarked on extensive electrification programs, with the realization that therein lie the best prospects of utilizing prime energy sources which are now and will be available to them in the future on a continuing basis.

So far as the U.S. is concerned, however, there will be no extended use of electric motive power until it can be demonstrated with irrefutable authority that fuel oil for use in self-contained motive power will not be economically available to the railroads as of a fairly definite point in future time. As of today, the oil-producing interests have demonstrated that this point in time is a comfortable span of years in the future. It is, moreover, quite evident that this point will not come without sufficient warning to permit of an orderly and financially bearable changeover to forms of motive power alternative to the diesel-electric.

Sidney Withington (retired) (New York, New Haven and Hartford Railroad, New Haven, Conn.): This paper suggests a number of controversial points in railroad operation which will doubtless bring out some vigorous discussion such as from the diesel-locomotive standpoint in construction, operation, and financing. The points of view of signal and communication engineers and of power companies should also be of interest in connection with problems raised by distribution of 60-cycle single-phase power in the trolley wire. It is to be hoped that contributions to the discussion will be received from operators of railroads in various parts of the world now using 50- or 60-cycle rail return power.

One of the serious limitations in the development of railroad electrification has been the cost of power especially in territory not served by hydroelectric plants. The authors of the paper compare energy costs of power with the cost of diesel fuel oil and have implied that the total or aggregate cost of power even at 60-per-cent load factor does not materially increase the cost of power delivered. Many of the power companies are very "liberal" in their method of calculating demand charges in terms of capital required per kilowatt of investment dedicated to railroad service. Furthermore, the tendency among utility companies is to shorten the integrated peak for billing purposes, which because of sharp demand swings and therefore of high short-time demands tends to increase the primary charge materially. Not long ago a 1-hour integrated demand was standard in railroad traction power billing but this has been recently modified in a number of cases.

The authors do not mention gas turbines. In spite of metallurgical problems which are yet ahead, it is probable that the gas turbines before long will be a deciding factor in power generation and that gas turbines will supersede diesel engines in locomotives. It may even be that an extended territory of electrified railroad may be economically served by gas-turbine stationary plants along the right-of-way, operated perhaps without attendance, cutting themselves on and off the line as required, thus saving high-voltage transmission!

H. F. Brown and **R. L. Kimball**: Mr. Webb's discussion is of great value, coming from a well-known power engineer very familiar with single-phase railway load problems and their effect on power-supply systems.

We agree with him that the scheme shown in Fig. 10 would not be looked upon with enthusiasm by any power engineers for a thickly populated territory having important industrial loads. It was designed for just one purpose, that of getting hydroelectric power to a railroad having no other domestic fuel source in an undeveloped territory, to build up the railroad traffic and to develop the territory, and to supply power to both at the minimum of investment. After the territory developed, it could afford to separate the traction load from the power load in a more conventional way.

It is agreed also that in Figs. 11, 12, and 13 the railroad loads will vary in each section and that some phase unbalance would be imposed on the power supply. As mentioned in the paper, some phase balancing equipment was included in the estimates shown in Fig. 15 because the power generating equipment could tolerate very little negative-sequence current. The phase-balancing scheme considered included series reactors between the supply generators and the railway load, with a shunt synchronous condenser at each end of the 66-kv line for regulation, each having a low X_2 value. For this particular study this was far cheaper than the 3-phase 2-phase scheme with switched capacitors and reactors at each supply point, with which the authors are familiar.

The scheme shown in Fig. 11 was based on the knowledge of the successful operation for many years of the single-phase load on the power supply mentioned in the last paragraph of Mr. Webb's commentary. The total railroad load in Fig. 11 is less than 60 per cent of the total load of the railroad referred to by Mr. Webb, and while the distribution of the load over the three phases as shown does not guarantee perfect phase balance at all times, the unbalance can never be more than 20 per cent of the single-phase railway load referred to by Mr. Webb as being all on one phase, which, as he says, "has operated satisfactorily to both parties" for many years.

Mr. Hamilton, while in agreement on the basic possibilities of commercial frequency electric traction, takes exception to:

1. The 20-year life assumed in the paper for the diesel-electric locomotive.
2. The omission, in the paper, of the necessity for including expensive signal changes in any electrification project.
3. The voltage used on the feeder in the 3-wire system shown in Figs. 9 through 13.
4. The taking of single-phase power at numerous points from alternate phases, instead of balancing the single-phase railroad load across all three phases at each supply point.

1. Regarding the expected life of the diesels: The authors are in agreement with Mr. Hamilton that the running gear (mechanical parts) and traction motors should have the same life expectancy as electric locomotives. This was implied in the paper by the statement that "in a major electrification" some of the diesels might be converted into electric locomotives.

Before adopting the 20-year life for the diesel (which the Interstate Commerce Commission [ICC] allows) for supporting the economic views expressed in the paper, the economics of the total investment and rising maintenance costs were carefully investigated for the following assumptions:

A. 20-year life of the entire locomotive.
B. 40-year life of unit, except engine; engine replaced, new, every 10 years.
C. 40-year life of unit, except engine; engine replaced, new, every $13^1/_3$ years.
D. 45-year life of unit, except engine; engine replaced, new, every 15 years.
E. Entire locomotive returned to builder for complete overhaul and rebuilding every $13^1/_3$ years, at 75 per cent of original cost, with "as new" guarantee.
F. Entire locomotive returned to builder for complete overhaul and rebuilding every 15 years, at 75 per cent of original cost, with "as new" guarantee.

The results of these investigations, assuming the same rising maintenance cost trends shown in Fig. 8, with straight-line depreciation, are shown in Table IV for comparison with each other and with electric locomotives with 40-year and 45-year life. 1,000-rail-hp units were assumed, making 120,000 miles per year, each costing $135,000, with the engine costing one-third of this amount or $45,000.

It will be seen from Table IV that the lowest cost assumption is the returning of the entire unit to the builder twice during the assumed life (if the builder will accept this arrangement twice). The next lowest is the complete renewal of the engine every 10 years. So far, neither of these methods has been adopted by the railroads as a standard policy. All other methods of maintenance, including that suggested by Mr. Hamilton (which is the most expensive), are within 3 per cent of the costs of the method assumed in the paper, viz., a complete new diesel unit every 20 years, which was adopted as conforming to allowed ICC practice. It is obvious from the tabulated results that the ICC statisticians who established 20 years as the fair life of the diesel locomotive knew their arithmetic (and their diesels).

2. The authors concur with Mr. Hamilton that some changes to signal systems are usually required when a line is electrified. These changes vary with the system of electrification installed and the signal system in use. With the 3-wire high-voltage commercial frequency system the current in the track should be much lower than with any other existing system. Signal-type bonding will be entirely adequate, and even one rail would suffice for the traction current in most cases. This opens up the possibilities of retaining much more of the existing signal equipment than heretofore, or of returning to simpler signal circuits. Certainly there will be no necessity of entirely abandoning existing roadside signal equipment. When signal changes are made for improved service, whether done at the time of electrification or later, their costs should be allocated to the improvement in service and not charged to electrification. Electrification, or any other improvement, can precipitate a number of other changes, but it should not necessarily be expected to bear all their costs. Such collateral charges are increased costs to the railroad, and must be taken into the electrification account, only if they constitute no improvement to the existing service or system, and then only for their unamortized value.

It is pointed out very clearly in the paper that line maintenance charges must be deducted from the difference in locomotive maintenance costs in comparing diesel and electric operating economies. There should be no greatly increased signal maintenance costs which Mr. Hamilton mentions. It would be just as proper to include hypothetical reductions in signal maintenance costs

Table IV

Assumption	40-Year Investment per 1,000-Rail-Hp, Dollars	Avg. Mntc. Cost per Mile over 40-Year Life, Cents	Combined Mntc. & Deprc. Cost per Mile, 40-Year-Life, Cents	Per Cent of Assumption A Unit Costs
A	270,000	19.5	25.2	100
B	270,000	18.0	23.7	94
C	225,000	19.7	24.4	97
D	225,000*	21.7*	25.9*	103
E	337,500	15.2	22.3	88.6
F	337,500*	16.3*	22.6*	89.5
Electric (40-year life)	135,000	13.0	15.8	62.5
Electric (45-year life)	135,000*	13.8	16.2*	64

* 45-year life assumed.

due to substitution of electric power for batteries, for example. None of these items are proper charges for or against electrification.

In connection with Mr. Hamilton's query regarding line maintenance costs: the figures given "per train mile" could in most cases be used as "per locomotive mile." The per-train-mile basis was used because in a number of cases the motive power is all multiple-unit motor cars, and no locomotives are used.

3. The 3-wire system used in the various diagrams in the paper has been used on most of the a-c electrified railroads in this country, and was designed primarily for inductive interference mitigation. The optimum voltage for the feeder usually should not be greater than twice the trolley-rail voltage, but it does not necessarily have to be twice, to obtain the best results. If the trolley-feeder voltage at 50 kv is high enough for the transmission of the maximum traction loads between the feed-in points, nothing is gained, and much is lost by going to a higher voltage.

The nominal voltage of 50 kv (which could also be 48 kv if 4 per cent made any great difference) was chosen for the trolley-feeder, with the rail midway between, after careful study, for the following reasons:

a. It is adequate for all normal traction loads, with normal catenary design, and with balancing transformer spacing normally required for the inductive co-ordination, and with the supply transformer spacing visualized for taps to various phases.

b. The insulation and the clearances are the same for the trolley and the feeder.

c. It allows closer spacing of trolley and feeder to each other and to the ground wire, which is also of fairly high conductivity and connected at frequent intervals to the rail.

d. This closer linkage of these three conductors tends, by their mutual inductance, to take the current out of the rail, which greatly reduces both the electromagnetic and the electrostatic induction.

e. It avoids the necessity of going to a higher insulation class and more expensive type for both the supply and balancing transformers and traction circuit breakers. (The 72-kv trolley-feeder proposed by Mr. Hamilton would necessitate a higher voltage rating for all this equipment.)

4. We have stated that the various schemes shown in Figs. 10–13 and Fig. 17 are a departure in some respects from concepts previously applied to traction circuits. In many details, however, they have already been used either on railroads or in commercial power applications. All have been carefully studied and discussed with manufacturers and power application engineers. All are feasible for application to the type of railroad service for which they are designed. Obviously Fig. 10 would not be applied to a densely populated busy commercial district, nor would a 4-track railroad be built to open an African jungle.

Mr. Hamilton is referred to our commentary on Mr. Webb's discussion, for his comments on Figs. 11 through 13.

Long distances between supply points are not necessarily desirable. More frequent taps from the various phases of the power network might possibly be more acceptable to the power supply system than a group of switched capacitors which Mr. Hamilton proposes as a standard. They would certainly be far cheaper. These points must be worked out with the power supply in all cases, and co-ordinated with acceptable phase-break locations on the railroad system. The balancing transformers on the 3-wire system of the railroad must be spaced as inductive co-ordination requirements dictate.

The authors are very familiar with the static phase-balancing scheme mentioned by Mr. Hamilton and Mr. Webb, and have purposely avoided such an expensive arrangement in the concepts outlined in this paper. Many phase-balancing schemes can be designed if needed. There is no reason, as yet, for considering any as absolutely required, or as the "standard basic arrangement" as Mr. Hamilton suggests. A great percentage of all the output of all the multitude of central stations is used for lighting. Every light is a single-phase load. These small loads are grouped into larger single-phase loads which are distributed and balanced across the various phases of the power networks until a well balanced 3-phase load on the generators results. Railroad traction loads are admittedly much larger, but there is no reason why they cannot be successfully distributed and balanced on large networks without too much expensive additional phase balancing apparatus. This is one of the major theses of this paper.

Dr. Huldschiner's discussion is of value in calling attention to locomotive design possibilities not mentioned in the paper. The authors are familiar with the various types of 50-cycle "direct" motors developed in Europe and which have been applied to the Hollenthal line in Germany, to the 50-cycle Annecy line in France, to the B.C.K. Railroad in the Belgian Congo, and for the French line under construction between Valenciennes and Thionville. They concur with Dr. Huldschiner that the possibilities of the a-c series commutator motor should not be overlooked.

The 50-cycle series commutator traction motor has been developed mainly by increasing the number of poles (and brushes), enlarging the commutator (as compared with the d-c motor), reducing the size of the motor by making it a "twin" or double armature motor, and by reducing its operating voltage. American manufacturers have contributed to this design through their European licensees, but have not built any of these motors since there are no 50-cycle power systems in this country.

To adapt this design for 60-cycle operation is possibly feasible, but it will entail some further expensive design changes and experimental research. The leading American manufacturers have felt that since the American railroads have almost universally adopted the 600-volt d-c traction motor, now used in the diesel-electric locomotives, the best economic solution for the adoption of commercial frequency electrification was to make this frequency operate motors of existing design through the rectifier. Consequently they have spent much time and effort in the development of the rectifier locomotive, which can operate on 25, 50, or 60 cycles as required.

As to weights and capacities: it is understood that one type of B_0-B_0 locomotive designed for the S.N.C.F. (French National Railways) with "direct" 50-cycle motors is rated at 2,500 hp. The transformer is designed to supply 10,000 amperes maximum in the secondary winding because the motors are designed for approximately 175 volts. The ignitron rectifier design uses the same B_0-B_0 running gear with standard 600-volt d-c motors, with correspondingly lower values of current in the transformer secondary winding. Its capacity is rated at 3,000 hp.

Series commutator 25-cycle motors have been built in this country for mounting on trucks similar to diesel-electric locomotive trucks, with dimensions not greatly different from d-c motors of the same rating. However, 60-cycle motors of these ratings have not yet been made or even designed. Possibly if and when a 60-cycle electrification is installed in this country, a stimulus will be given to designing such a motor. The present indications are, however, that the maintenance cost of such a motor, as well as the first cost, might be much greater than the combined similar costs of the standard d-c motor and the rectifier.

The power factor of such motors could be lower than that of the rectifier locomotive. It probably would not be much better in this respect. If it should prove to have economic possibilities it will eventually be developed.

The authors are more or less in agreement with Dr. Huldschiner in the evaluation of the motor-generator locomotive. Its principal possibilities are only for power-factor improvement in the event that the rectifier locomotive cannot be designed for 60 cycles with a good power factor. Present indications, based on 25-cycle ignitron rectifier locomotives in use, indicate a power factor which should be acceptable to the power supply companies.

Mr. Perkinson's comments are appreciated as coming from one who is a recognized authority in this field. The authors are gratified that he seems to be in general agreement with them regarding the fundamental economic superiority of electric operation as so clearly and concisely summed up in the final six paragraphs of his discussion. The main points wherein he differs from the opinions expressed in the paper appear to be in the evaluation of maintenance costs, power costs, and depreciation costs.

Relative to maintenance costs, he refers to these as becoming "stabilized." A major thesis of the paper is that maintenance costs of all types of motive power, whether steam, diesel, or electric, are constantly rising over their useful life. This is a fact and not theory, as shown in Figs. 5 through 7(B). Further evidence of this is submitted in Fig. 18, which is taken from data originally published in 1932 by Thomas R. Cook of the Baldwin Locomotive Works.[1] In this figure, 1932 costs have been equated to 1952 values, and the diesel and electric maintenance costs shown in Figs. 5, 7(A), and 7(B) have been added for comparison. There is no evidence to indicate that diesel-electric costs will not follow the same trends, possibly at a different rate, which have been determined during the past 45 years by both steam and electric motive power.

Another recognized authority, K. Cartwright, in a paper presented before the Pan American Railway Congress in June 1953, entitled "Factors Relating to Selection of Type of Locomotive for Various Operations, (Steam, Diesel-electric, Electric)," states:

"The maintenance cost of any new loco-

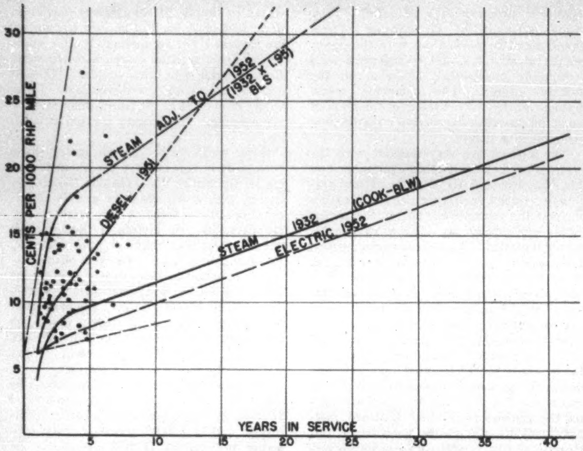


Fig. 18. Rise in locomotive maintenance costs with age

motive is low during the first year of operation. Costs then begin to rise and follow a rising trend throughout the life of the locomotive. This cost trend has been worked out for steam locomotives in the U.S.A. for a period of 40 years. It indicates that taking the maintenance cost for the first year as 100%, the cost at a ten year age is approximately 200%; 20 yr. age, 255%; 30 yr. age, 310%. . .

"Sufficient reliable data have not been accumulated to establish similar costs for diesel-electric or electric locomotives. It is manifest, however, that the repair costs of these units will advance with age, and judging from such costs as are available it appears that on the basis of the cost per horsepower unit, the maintenance cost of the diesel-electric locomotive will be comparable with the steam and that of the ignitron rectifier type of electric locomotive, appreciably less."

It is therefore apparent that the basic maintenance costs Mr. Perkinson has used in the beginning of his commentary for electric locomotives are too low, for they disregard entirely the probable rising trend over the life. To double these costs to obtain diesel costs may check with the present diesel costs, but will also neglect to include the rising trends in these costs which the paper endeavors to evaluate. Thus the total costs he derives for comparison with those shown in Table III are much too low, and his contention that the costs shown in the paper are too high is invalid. The costs shown in Table III are the calculated average costs over the life, based on present trend indications.

Relative to his difference of opinion on power-fuel costs: In the example shown in Table III electric energy is assumed to be purchased for 8 mills at the point of supply to the railroad, and the use of this energy was assumed to be at the rate of 35 watt-hours per gross ton mile. This rate is more than ample to include all of the transmission and conversion losses mentioned by Mr. Perkinson, which in themselves are not greatly different in magnitude from the engine and generator conversion losses on the diesel. Both locomotives working at the same speed must use the same number of watt-hours per ton mile at the rim of the driving wheels. The assumptions are that power can be purchased for 8 mills, and that diesel fuel will cost 10 cents per gallon, which give the comparative costs shown in Table III. This is still somewhat favorable to the diesel, as fuel handling costs by the railroad, which may add greatly to the cost of fuel on board the locomotive, are not included.

The 8-mill power rate assumed was also very carefully checked as to its possibility of being realized, in one or two districts where power is available and general railway electrification appears to be an economic possibility, and was found to be well within the range of present-day rate values at the load factor assumed, even with all steam generation.

Mr. Perkinson states that "the matter of proper depreciation rates for use with the two types of motive power is one that is open to debate." Obviously the capital investment in the motive power involved must be written off during its life. If two-thirds of the diesel locomotive is good for 45 years, and a new engine is purchased every 15 years then the investment in one complete locomotive and two engines must be written off in 45 years. If a locomotive is completely rebuilt in 15 years for 75 per cent of its original cost, then 75 per cent of its original cost must be written off in 15 years. This is elementary accounting. Table IV clearly indicates how little difference there is in the combined maintenance and depreciation charges between Mr. Hamilton's or Mr. Perkinson's assumptions and those made in our paper. The debate goes right back to maintenance costs, and the necessity for keeping these costs in a manner which will show when retirements should be made for the greatest economy.

The authors agree also that a number of "electric locomotives have gone to the scrap pile after 25 to 30 years of life because they have been superseded by more economical types. . ." Mr. Perkinson doubtless knows that most of these have been motor-generator type locomotives.

To state that diesel "locomotives will not wear out because in the normal maintenance operations they are being constantly renewed, and in some cases, improved with new or repaired components" is hardly compatible with "stabilized" maintenance costs, unless these are high enough to absorb all the depreciation charges. If the statement were true, it would apply equally to electric locomotives. However, in neither case is it factual. Everything that does work eventually wears out to the point where it is too expensive to maintain further.

In his concluding paragraph, Mr. Perkinson believes that the economic availability of fuel oil alone will decide the time when the diesel will be replaced by the electric locomotive. It is our opinion that the major deciding factor will be rising maintenance costs, and by the time this trend is clearly established, the fuel differential, already apparent, will give additional economic weight to the change.

Mr. Withington comments on the cost of power having been a limitation to the development of electrification in the past. This may have been true in certain areas where the power supply was limited and where general railway electrification might otherwise have been feasible. This situation, however, has been greatly changed within the past decade. There are large power networks today served mainly by steam stations which not only could, but would if called upon, assume sizeable general railway loads, visualized as having load factors of 60 to 75 per cent.

Some existing power contracts based on 40-per-cent load factor could show a 2-mill reduction in the cost per kw-hr if at a load factor of 65 per cent. High half-hourly or hourly peaks which are characteristic of many present electric railway operations for passenger terminals and suburban traffic, and which determine many of the "demand" charges, are not visualized in many of these future possible general railway electrifications.

All charges making up the usual cost of electric power have been taken into consideration in the costs assumed in Table III. The only place "energy" costs alone were considered was in the case where diesel locomotives might be operated into electrified territory.

The gas-turbine locomotive and the coal-burning steam-turbine locomotive were not mentioned in the paper as being somewhat outside its scope. If the gas turbine ever should become a serious competitor of the diesel locomotive, the electric would have to compete with those factors which would be superior to the diesel. So far, such factors have not become too apparent, nor do there seem to be any technological developments of any real promise under way to indicate that such will be developed within the next decade. Such factors would have to include a longer life and a lower maintenance cost than the diesel.

As to a multiplicity of gas turbines stationed along an electrified railroad, operated without attendance and eliminating high-voltage transmission, this would not be able because of high maintenance costs to compete with the growing network of existing transmission systems which are crossing and recrossing the railroads in many areas today. In districts not served by any power transmission networks, such isolated plants could not possibly compete with the present mobile diesel-electric locomotive.

The future railway electrification envisioned no longer requires railroad-owned power stations, transmission lines, converter stations, or substations of any kind. It takes full advantage of the existing investment in these large power networks and facilities just as the other large industrial users of electric power in these areas are already doing. This places railway electrification in an entirely new position in so far as railroad investment is concerned.

Electric power is now available in nearly all areas where the railroad traffic is important enough to consider its use. This now places these railroads in a position to realize on the operating economies of electrification at any time in the future.

In closing, we wish to thank those who have contributed to the value of this paper by their discussion. It was not intended in this paper to disparage the remarkable service the diesel-electric locomotive has rendered to the railroads of this country during the past decade. This type of motive power will in all probability continue in service for many years to come, either as the sole type, on certain railroads where electrification cannot be justified, or as a useful and economic adjunct to those railroads which ultimately will again turn to electrification for its greater economy.

If the paper reawakens interest in a subject that has been allowed to become dormant, or if it calls attention to operating cost factors which should possibly receive a little more attention, its purpose will have been accomplished.

REFERENCE

1. HOW AGE AFFECTS LOCOMOTIVE MAINTENANCE COSTS, Thomas R. Cook. *Railway Mechanical Engineer*, New York, N. Y., Sept. 1932, p. 347.

Contactor Servomechanisms Employing Sampled Data

C. K. CHOW
ASSOCIATE MEMBER AIEE

Synopsis: This paper introduces a sinusoidal response method of analysis and synthesis of a contactor servomechanism employing sampled data. Because of the presence of sampler, clamper, and contactor, the system is discontinuous in time as well as being nonlinear. The sinusoidal method provides a useful guide in analysis and synthesis. It enables one to determine the conditions of occurrence of self-sustained oscillations and their amplitudes and frequencies; and to evaluate critical values of certain system parameters. Electronic analogue computer techniques are employed to check the accuracy of the sinusoidal response method.

THE particular nonlinear servo system to be investigated in this paper is a contactor servomechanism employing sampled data in the error channel. It is a combination of sampled data (sampling) servomechanism, and contactor (or relay) servomechanism. Because of the nonlinear and discontinuous characteristics of the system, an exact solution of system performance can be obtained only by considerably laborious computation. However, much can be learned about the system by employing a sinusoidal response technique. The sinusoidal response method has been applied to contactor servomechanisms, notably by Goldfarb[1] and Kochenburger,[2] to pulsed servos,[3] and to certain other nonlinear automatic control systems.[4-7]

The "critical point" on the complex plane for a contactor servomechanism is a curve which is amplitude-dependent and frequency-invariant.[1-3] In some cases, as pointed out by Johnson,[4] it may be a family of amplitude-variant loci, one locus being constructed for each frequency. Radically different from the results obtained by these investigators, the critical point for the servo system treated in this paper is a family of regions on the complex plane, one region for each dimensionless period of integer value.

List of Symbols

DIMENSIONAL NOTATIONS

a_i = time constant in the numerator of $KG(p)$
b_j = time constant in the denominator of $KG(p)$
$G(p)$ = frequency variant part of $KG(p)$
K = frequency invariant part of $KG(p)$, the over-all forward loop gain
$KG(p)$ = transfer function of the linear part of the forward loop
p = the complex variable of Laplace transform
t = elapsed time, seconds
T_s = sampling period, seconds; used as the time base for establishing dimensionless notations
$\theta_b = KT_s^{\lambda}$ = output base for establishing dimensionless notations
θ_c = Clamper output
$\theta_e = \theta_i - \theta_0$ = system error
θ_i = system input
θ_o = system output
θ_r = contactor output
θ_Δ = inactive zone of the contactor

DIMENSIONLESS NOTATIONS

$c = \theta_o/\theta_b$ = system output
$C(s)$ = Laplace transform of $c(\tau)$
D = derivative operator with respect to τ, $= d/d\tau$
e = base of natural logarithm
$g(s)$ = transfer function of the linear part of the forward loop
i = running index
$j = \sqrt{-1}$
k = constant, = 0, $^1/_2$, or 1
m = clamper output
n = running index
N = describing function of the nonlinear element
$-\dfrac{1}{N}$ = generating phasor of the critical region
$-\dfrac{1}{N}\Big|_{\min}$ = generating phasor of the boundary of the critical region
$-\dfrac{1}{N}\Big|_{\max}$ = generating phasor of the boundary of the critical region
Q = integer
$r = \theta_i/\theta_b$ = system input
$R(s)$ = Laplace transform of $r(\tau)$
$s = T_s p$ = complex variable of Laplace transform
u = angular frequency
$y = f(x)$ = a given function of x
Y_n = Fourier coefficient of y
$\alpha_i = a_i/T_s$ = time constant in the numerator of $g(s)$
$\beta_j = b_j/T_s$ = time constant in the denominator of $g(s)$
γ = magnitude of Γ_1
Γ_1 = fundamental frequency component of $\Gamma(\tau)$
$\Gamma = \theta_r/\theta_b$, contactor output
δ = duration time of each contactor corrective signal
$\Delta = \theta_\Delta/\theta_b$ = inactive zone of the contactor
Δ_c = critical value of Δ
$\epsilon = \theta_e/\theta_b$ = system error
ϵ_s = sampled system error, sampler output

Paper 54-123, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 20, 1953; made available for printing December 18, 1953.

C. K. CHOW is with The Pennsylvania State University, State College, Pa.

The material presented in this paper is based on the thesis submitted by the author in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Cornell University. Acknowledgment is due to Prof. Wilbur E. Meserve for supervising the thesis and reviewing the paper. Also acknowledged is the invaluable assistance given by Dr. David A. Kahn and William K. Kindle of the Cornell Aeronautical Laboratory, Inc., Buffalo, N. Y., in conducting the studies of the simulated servo systems on a REAC computer.

ζ = damping ratio
$\theta = (2\pi/T)$ radians
λ = order of integration factor
μ = constant of integer value
ν = constant of integer value
$\tau = t/T_s$ = time
τ_a, τ_b = contactor switching instants
T = period
ϕ = phase angle between ϵ and the sampling instants
ψ = phase angle of $-1/N$

Functional Description of the System

The block diagram of the particular servomechanism under investigation is shown in Fig. 1. It is assumed that all parts of the servo system except the sampler, clamper, and contactor are linear, and the sampling rate is uniform.

Both system input $\theta_i(t)$ and output $\theta_o(t)$ are continuous functions of time t. The sampler closes momentarily at discrete time intervals, that is, at $t=o$, T_s, $2T_s$, $3T_s$, . . ., where T_s is the constant sampling period. Therefore error detection is accomplished only at discrete instants, and a signal is fed intermittently to the clamper at $t=o$, T_s, $2T_s$, $3T_s$ . . ., the sampling constants. No data at all are supplied in the intervals separating these sampling instants. The output of the sampler consists of a train of equally spaced pulses of short duration whose envelope is the control error function $\theta_e(t)$.

The clamper or holding circuit receives sampled signals from the sampler, and delivers at any given instant the same signal as was obtained from the last data received. The output of the clamper is a step function of time which changes its magnitude abruptly at the sampling instants. The clamper output actuates the contactor.

A contactor is essentially a discontinuous power amplifier. The peculiar characteristic of the contactor is its discontinuous action; that is, it is capable of applying only one of a limited number of discrete magnitudes of corrective signal. In other words, the output signal of the contactor θ_r is a discontinuous function of input to the contactor. In this paper only a simple form of contactor having an operating characteristic, as shown in Fig. 2, is considered.

Mathematically the contactor output, θ_r can be written as a function of contactor input θ_c in the following form

$$\theta_r(\theta_c) = \begin{cases} +1 & \text{for } \theta_c > \theta_\Delta \\ 0 & \theta_\Delta > \theta_c > -\theta_\Delta \\ -1 & -\theta_\Delta > \theta_c \end{cases} \tag{1}$$

where θ_Δ is the inactive zone of the contactor, which is expressed in the units of output quantity, for example in degrees, if the output quantity θ_o is measured in degrees.

The contactor is assumed to respond instantaneously to the control signals, as in many cases the dynamic lag involved may be negligible. In cases where the dynamic lag cannot be neglected, it can be approximately represented by a finite time delay in seconds; the sinusoidal method is still applicable.

The output of the contactor actuates the linear part of the forward loop which can be represented by a transfer function of the following form

$$KG(p) = \frac{K \prod_{i=1}^{\nu} (1+a_i p)}{p^\lambda \prod_{j=1}^{\mu} (1+b_j p)} \tag{2}$$

where $\lambda+\mu>\nu$. a_i's, b_j's are time constants, either real or complex, and $a_i \neq b_j$ for all i and j; K is the forward loop gain constant, and p is the complex variable of Laplace transform. The output of the system is then fed back and completes the servo loop.

Because of the existence of the sampler, clamper, and contactor, the system is quantized both in time and magnitude, and is therefore discontinuous with respect to time, and is nonlinear.

Dimensionless Representation

Nondimensionalized quantities are used throughout this investigation. For convenience T_s and KT_s^λ are chosen respectively as time and signal basic units. On these bases, the dimensionless time τ is obtained by dividing the actual time t in seconds by T_s that is

$$\tau = t/T_s \tag{3}$$

On the dimensionless base the sampler will close momentarily at $\tau = 0, 1, 2, 3 \ldots$ integer values of τ.

By dividing the base KT_s^λ all the system signals are converted into the dimensionless form. The dimensionless system input r output c, and inactive zone of the contactor Δ are related to the actual quantities by the equation

$$r = \frac{\theta_i}{KT_s^\lambda} \tag{4}$$

$$c = \frac{\theta_o}{KT_s^\lambda} \tag{5}$$

$$\Delta = \frac{\theta_\Delta}{KT_s^\lambda} \tag{6}$$

The dimensionless form of the transfer function $KG(p)$ is, then

$$g(s) = \frac{\prod_{i=1}^{\nu} (1+\alpha_i s)}{s^\lambda \prod_{j=1}^{\mu} (1+\beta_j s)} \tag{7}$$

where

$\alpha_i = a_i/T_s$

$\beta_j = b_j/T_s$

$s = T_s p$

Fundamental Equations

Consideration of the nature of components of the system furnishes the following basic dimensionless equations which completely describe the system behavior

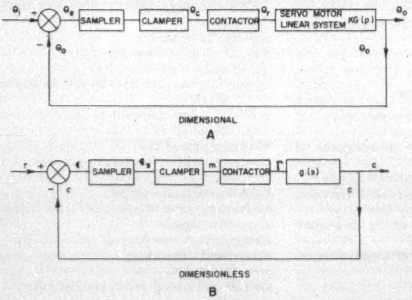


Fig. 1 (left). Block diagram of the servomechanism

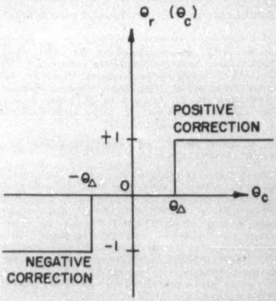


Fig. 2 (right). The operating characteristic of the contactor

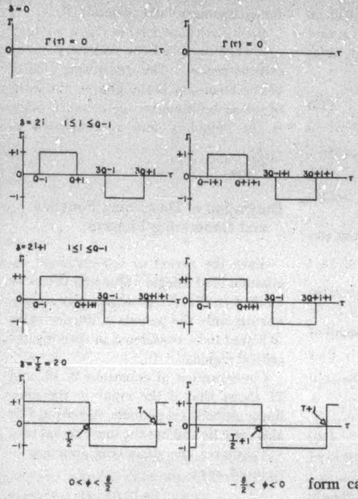


Fig. 3 (left). Periodic contactor output of period T, where T=4Q

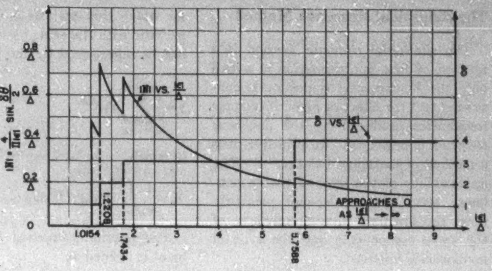


Fig. 4 (above). Plot of the describing function N of T=8 and $\phi=\pi/18$ together with the corresponding value of δ

$$\epsilon(\tau)=r(\tau)-c(\tau) \tag{8}$$

$$\begin{aligned}\epsilon_s(\tau)&=\epsilon(\tau) \text{ for } \tau=0, 1, 2, 3 \ldots \text{ integers}\\ &=0 \text{ elsewhere}\end{aligned} \tag{9}$$

$$m(\tau)=\epsilon_s(n) \quad \text{for } n\leq\tau<n+1 \tag{10}$$

where

$n=0, 1, 2, 3 \ldots$ integers

$$\Gamma(m(\tau))=\begin{cases}+1 & \text{for } m(\tau)>\Delta\\ -1 & m(\tau)<-\Delta\\ 0 & -\Delta<m(\tau)<\Delta\end{cases} \tag{11}$$

$$\frac{1}{g(D)}c(\tau)=\Gamma(\tau) \tag{12}$$

where $D=d/d\tau$ is a dimensionless time derivative operator and ϵ, ϵ_s, m and Γ represent respectively the dimensionless system error, sampler output, clamper output, and contactor output.

Equation 12 is a piecewise linear differential equation with constant coefficients, having a forcing function, $\Gamma(\tau)$, which is an unknown step function of time τ and dependent nonlinearly on $\epsilon(\tau)=r(\tau)-c(\tau)$.

For any arbitrary given system input $r(\tau)$ the system response $c(\tau)$ can be uniquely determined by the classical method of solving the foregoing equations and matching the end conditions at the discontinuities of contactor output Γ. It can also be determined by a semigraphical method employing Laplace transforms.[8] This semigraphical method has been applied successfully to relay servomechanisms by Kahn.[9] In general, owing to the discontinuity and nonlinearity, no analytic solution in a closed form can be obtained. Since the main purpose of this paper is to demonstrate the application of the sinusoidal response method to this particular system, no further consideration of these exact methods will be given.

Periodic Solution

In this section some pertinent facts concerning the periodic oscillation are derived. If $c(\tau)$ is a periodic function of time having a period T, then the quantity $(1/g(D))c(\tau)$ will also be a periodic function of τ having the same period T. By virtue of equation 12, this means that $\Gamma(\tau)$, the contactor output, is a periodic function of τ having the same period T.

The contactor output consists of nothing but a train of positive, negative rectangular pulses and zero values, and it changes its values only at the integer values of τ. Consequently, the durations of positive and negative pulses, also of zero corrections are integers.

Let τ_a be one of the contactor switching instants, that is, the instants at which $\Gamma(\tau)$ changes its values. By the periodicity of $\Gamma(\tau)$, at T later $\tau=\tau_a+T$, is also a switching instant, called t_b.

Therefore

$$T=\tau_b-\tau_a \tag{13}$$

Since both τ_a and τ_b are integers, their difference T is an integer. This means that the period of periodic system response is an integer.

For the case wherein the servo system has at least one integration factor, i.e., $\lambda\geq1$ in the expression of $g(s)$, equation 7, if $c(\tau)$ is periodic and expressible in Fourier series, then by the aid of equation 12, $\lambda\geq1$ implies that the average value of periodic Γ is zero. In other words, in the periodic response the duration of positive contactor correction signal is equal to that of negative contactor correction signal during one period. This duration time is designated as δ, and is used to characterize the mode of self-sustained oscillation. For detailed proof, see reference 8. For simplicity, only the systems for which $\lambda\geq1$ are considered in this paper. The critical regions are derived under this assumption.

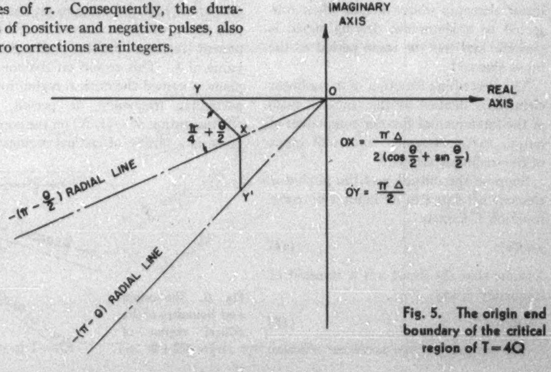


Fig. 5. The origin end boundary of the critical region of T=4Q

The Sinusoidal Response Method

The sinusoidal response method is a method of first approximation in which only the fundamental frequency is considered and the effect of higher harmonic frequencies produced by the nonlinear elements is neglected. In principle, it is a method of harmonic balance.[10,11] The applicability of the method to the nonlinear servo system depends primarily upon the low-pass filtering property of the servo components and the type of nonlinearity involved.

Although being an approximate method, the sinusoidal response method provides a useful guide in the analysis and synthesis. It enables one to determine the conditions of occurrence of self-sustained oscillation of the nonlinear system and their frequencies and amplitudes. For a given synthesis problem, the sinusoidal method makes possible the determination of the critical value of inactive zone and the selection of means for reducing the critical values of Δ to an optimum value.

The value of the inactive zone, necessary to just prevent self-sustained oscillation in the system, is defined as the critical value of Δ, designated as Δ_c. If for a given system self-sustained oscillation is permissible, then the actual value of inactive zone may be selected less than Δ_c. Otherwise, the actual value of Δ to be used must be somewhat greater than Δ_c.

In the sinusoidal response method the behavior of the nonlinear element is represented by the describing function. This representation is based upon the assumption that the input signal to the nonlinear element is sinusoidal and that the resulting output can be satisfactorily represented by its fundamental Fourier component so far as the performance of the whole system is concerned. This method is applicable only to those nonlinear elements whose output when subjected to a sinusoidal driving signal is periodic and has the same period as the input sinusoid.

The describing function of a nonlinear element is defined as the complex ratio of the fundamental Fourier component of output to the assumed sinusoidal input of the nonlinear element.

Suppose the output y of the nonlinear element is related to its input x by some function f, namely

$$y=f(x) \tag{14}$$

Assume that the input x is a sinusoid of frequency u, say

$$x=\mathrm{Re}Xe^{ju\tau} \tag{15}$$

The output y of the nonlinear element for which this analysis is useful will be periodic with the same period $2\pi/u$, and can be represented by the Fourier series

$$y=\frac{1}{2}\sum_{n=-\infty}^{\infty} Y_n e^{jnu\tau} \tag{16}$$

in which

$$Y_n=\frac{1}{\pi}\int_0^{2\pi} f(x)e^{-jun\tau}du\tau \tag{17}$$

The describing function N for the element is defined as

$$N=\frac{Y_1}{X} \tag{18}$$

The fundamental frequency component of x is, then

$$\mathrm{Re}\,Y_1e^{ju\tau}=\mathrm{Re}NXe^{ju\tau}$$

If the input sinusoid is

$$x=|\epsilon|\sin u\tau \tag{19}$$

then equation 18 reduces to

$$N=\frac{1}{\pi|\epsilon|}\int_0^{2\pi} f(|\epsilon|\sin u\tau)\sin u\tau du\tau + \frac{j}{\pi|\epsilon|}\int_0^{2\pi} f(|\epsilon|\sin u\tau)\cos u\tau du\tau \tag{20}$$

The describing function N, consequently its negative reciprocal $-1/N$, is a complex variable which can be represented by a phasor in a complex plane. On the complex plane a phasor originating from the origin of the plane, and terminating at the point which has a coordinate of $\mathrm{Re}(-1/N)+j\mathrm{Im}(-1/N)$ is called the generating phasor, also designated by $-(1/N)$. In the system under consideration, the sampler, clamper, and contactor in cascade are lumped together and considered as a single nonlinear element.

For a given period of integer value, the locus of the terminal of the generating phasor $-(1/N)$ is a region on the complex plane, which consists of several component regions, one for each possible value of δ. This region on the complex plane is named the critical region for the particular frequency or period. The representation of $-(1/N)$ on the complex plane is a family of critical regions, one for each integer value of period T.

Each critical region has its own boundary which is called the boundary of the critical region. The generating phasor of the boundary is the phasor, the locus of whose terminal on the complex plane is the boundary and is designated as $-\left.\frac{1}{N}\right|_{\min}$ and $-\left.\frac{1}{N}\right|_{\max}$.

Derivation of Describing Function and Generating Phasors

Since the period of self-sustained oscillation is of integer value and it can be easily seen that it cannot be unity, consequently only the periods of integer value $\geqslant 2$ need to be considered in deriving the critical regions.

Consideration of equations 9, 10, and 11 shows that if the input to the nonlinear element ϵ is periodic, its output Γ is also periodic and has the same period as ϵ.

To derive the describing function, a sinusoid, say

$$\epsilon=|\epsilon|\sin(u\tau+\phi)=|\epsilon|\mathrm{Im}e^{j(u\tau+\phi)} \tag{21}$$

is taken as the input to the nonlinear element where $u=2\pi/T$. T is an integer $\geqslant 2$.

For the purpose of deriving the describing function and critical region, the value of ϕ needed to be considered ranges from $-(\theta/2)$ to $\theta/2$, where $\theta=2\pi/T$ because for any value of ϕ outside this range a proper shift in τ-axis by an integer value will bring the value of ϕ into the range mentioned. Such a shift in τ-axis by integer value will in no way alter the relation of output Γ with input ϵ.

The general expression of the describing function and the generating phasors are derived for any given period of integer value T. The derivation is divided into four cases according to the value of T as shown in the following.

CASE 1

$T=4Q$ where Q is a positive integer. This case includes $T=4, 8, 12\ldots$

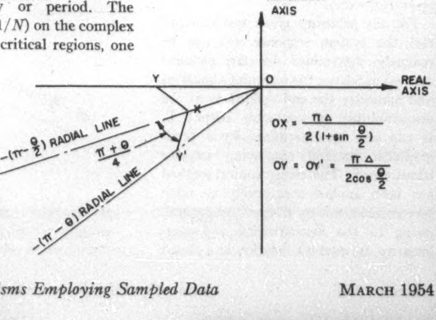

IMAGINARY AXIS
REAL AXIS
O
Y
X
Y'
$-(\pi-\frac{\theta}{2})$ RADIAL LINE
$\frac{\pi+\theta}{4}$
$-(\pi-\theta)$ RADIAL LINE
$OX=\frac{\pi\Delta}{2(1+\sin\frac{\theta}{2})}$
$OY=OY'=\frac{\pi\Delta}{2\cos\frac{\theta}{2}}$

Fig. 6. The origin and boundary of the critical region of T=4Q+2

$\epsilon(\tau)$ is sampled at $\tau=0, 1, 2, \ldots$ resulting in a train of sampled data at sampler output ϵ_s whose value at $\tau=i$ is

$$\epsilon_i=\epsilon_s(\tau)\big|_{\tau=i}=|\epsilon|\sin(i\theta+\phi) \tag{22}$$

where $i=0, 1, 2,\ldots$ a non-negative integer.

Obviously

$$\epsilon_{i+T}=\epsilon_i \tag{23}$$

$$\epsilon_{i+\frac{T}{2}}=\epsilon_i$$

For a given Q it is almost trivial that for

$$\theta<|\phi|<\frac{\theta}{2}$$

$$\begin{aligned}&\sin(Q\theta+|\phi|)>\sin[(Q-1)\theta+|\phi|]\\&>\sin[(Q+1)\theta+|\phi|]\\&>\ldots\\&>\sin[(Q-i)\theta+|\phi|]>\sin[(Q+i)\theta+|\phi|]\\&>\ldots\\&>\sin(Q\theta+|\phi|)>\sin[(2Q-1)\theta+|\phi|]\\&>\sin|\phi|>0\end{aligned} \tag{24}$$

The relations among the values of periodic sampled data are obtained from equations 22 and 24 as follows:

For

$$0<\phi<\frac{\theta}{2}$$

$$\begin{aligned}&\epsilon_Q>\epsilon_{Q-1}>\epsilon_{Q+1}>\epsilon_{Q-2}>\epsilon_{Q+2}\\&>\ldots>\epsilon_{Q-i}>\epsilon_{Q+i}\\&>\ldots>\epsilon_{2Q-1}>\epsilon_0>0\end{aligned} \tag{25}$$

For

$$0>\phi>-\frac{\theta}{2}$$

$$\begin{aligned}&\epsilon_Q>\epsilon_{Q+1}>\epsilon_{Q-1}>\epsilon_{Q+2}>\epsilon_{Q-2}\\&>\ldots>\epsilon_{Q+i}>\epsilon_{Q-i}\\&>\ldots>\epsilon_{2Q}=-\epsilon_0>0\end{aligned} \tag{26}$$

The value of δ may be $0,1,2,\ldots n\ldots$ up to $T/2$. To have $\delta=n$, it requires that n of $\epsilon_0,\epsilon_1\ldots\epsilon_{T/2-1}$ exceed the value Δ, the inactive zone of the contactor while the other $(T/2)-n$ ϵ_s's do not. By the aid of equation 24, this requirement imposes the following conditions on ϵ to ensure each value of δ

$$\delta=0\rightleftharpoons|\epsilon|\sin\left(\frac{\pi}{2}+|\phi|\right)<\Delta$$

that is, $|\epsilon|\cos\phi<\Delta$

$\delta=2i$, where $1\leq i\leq Q-1$

$$\rightleftharpoons|\epsilon|\sin[(Q-i)\theta+|\phi|]>\Delta>|\epsilon|\sin[(Q+i)\theta+|\phi|]$$

$\delta=2i+1$, where $0\leq i\leq Q-1$

$$\rightleftharpoons|\epsilon|\sin[(Q+i)\theta+|\phi|)>\Delta>|\epsilon|\sin(Q-i-1)\theta+|\phi|]$$

$$\delta=T/2\rightleftharpoons|\epsilon|\sin|\phi|>\Delta \tag{27}$$

Equation 27 gives the limiting values of $|\epsilon|$ for each value of δ in terms of Δ and ϕ as follows

$\delta=0$

$$|\epsilon|_{\min}=0$$

$$|\epsilon|_{\max}=\frac{\Delta}{\cos\phi} \tag{28A}$$

$\delta=2i$ where $1\leq i\leq Q-1$

$$|\epsilon|_{\min}=\frac{\Delta}{\sin[(Q-i)\theta+|\phi|]}=\frac{\Delta}{\cos(iQ-|\phi|)}=\frac{\Delta}{\cos\left(\frac{\delta}{2}\theta-|\phi|\right)}$$

$$|\epsilon|_{\max}=\frac{\Delta}{\sin[(Q+i)\theta+|\phi|]}=\frac{\Delta}{\cos(iQ+|\phi|)}=\frac{\Delta}{\cos\left(\frac{\delta}{2}\theta+|\phi|\right)}$$

$\delta=2i+1, 0\leq i\leq Q-1$ (28B)

$$|\epsilon|_{\min}=\frac{\Delta}{\sin[(Q+i)\theta+|\phi|]}=\frac{\Delta}{\sin(iQ+|\phi|)}=\frac{\Delta}{\sin\left(\frac{\delta-1}{2}\theta+|\phi|\right)}$$

$$|\epsilon|_{\max}=\frac{\Delta}{\sin[(Q-i-1)\theta+|\phi|]}=\frac{\Delta}{\cos[(i+1)\theta-|\phi|]}=\frac{\Delta}{\cos\left(\frac{\delta+1}{2}\theta-|\phi|\right)}$$

$\delta=T/2$

$$|\epsilon|_{\min}=\frac{\Delta}{\sin|\phi|}$$

$$|\epsilon|_{\max}=\infty$$

Consideration of the nature of nonlinear elements, equations 22 through 26, and conditions of equation 28 enables one to construct the contactor output $\Gamma(\tau)$ which is given in Fig. 3.

After being analyzed into Fourier series, the fundamental frequency component of $\Gamma(\tau)$ is found to be $\Gamma_1(\tau)$ as follows

$\delta=0$ $\qquad$ $\Gamma_1(\tau)=0$

$\delta=2i$ $\qquad$ where $1\leq i\leq Q$

$$\Gamma_1(\tau)=\begin{cases}\gamma\,\mathrm{Im}\,e^{ju\tau} & 0<\phi<\frac{\theta}{2}\\ \gamma\,\mathrm{Im}\,e^{j(u\tau-\theta)} & 0>\phi>-\frac{\theta}{2}\end{cases}$$

$\delta=2i+1$ $\qquad$ $0\leq i\leq Q-1$

$$\Gamma_1(\tau)=\gamma\,\mathrm{Im}\,e^{j\left(u\tau-\frac{\theta}{2}\right)}\qquad 0<|\phi|<\frac{\theta}{2} \tag{29}$$

where

$\gamma=\frac{4}{\pi}\sin\frac{\delta\theta}{2}$ is the amplitude of $\Gamma_1(\tau)$

By definition, the describing function is

---

given by the phasor equation

$$\frac{\Gamma_1}{\epsilon}=\frac{\Gamma_1}{|\epsilon|e^{j(u\tau+\phi)}}$$

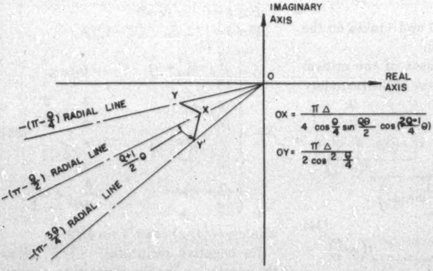


Fig. 7. The origin end boundary of the critical region of T=4Q+1

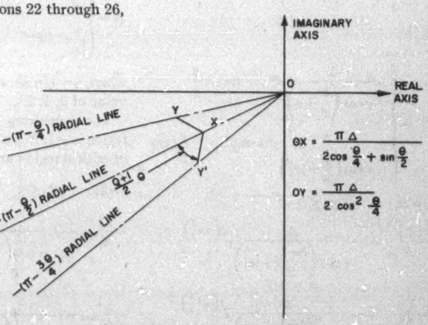


Fig. 8. The origin end boundary of the critical region of T=4Q+3

as

$\delta=0 \qquad N=0$

$\delta=2i \qquad 1\le i\le Q$

$$N=\begin{cases}\dfrac{\gamma}{|\epsilon|}e^{-j\phi} & 0<\phi<\dfrac{\theta}{2}\\[2ex] \dfrac{\gamma}{|\epsilon|}e^{-j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}\tag{30}$$

$\delta=2i+1 \quad 0\le i\le Q-1$

$$N=\frac{\gamma}{|\epsilon|}e^{-j\left(\phi+\frac{\theta}{2}\right)} \qquad 0<|\phi|<\frac{\theta}{2}$$

Since N is zero for $\delta=0$, the component critical region of $\delta=0$ is at infinity of the complex plane. In other words, there is no critical region for $\delta=0$, and no self-sustained oscillation characterized by $\delta=0$ will occur in the actual system. This is also intuitively obvious.

The generating phasor of the critical regions is $-1/N$ as given in equation 31

$\delta=2i \qquad 1\le i\le Q$

$$-\frac{1}{N}=\begin{cases}-\dfrac{|\epsilon|}{\gamma}e^{j\phi} & 0<\phi<\dfrac{\theta}{2}\\[2ex] -\dfrac{|\epsilon|}{\gamma}e^{j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}$$

$\delta=2i+1 \qquad 0\le i\le Q-1$

$$-\frac{1}{N}=-\frac{|\epsilon|}{\gamma}e^{j\left(\phi+\frac{\theta}{2}\right)} \qquad 0<|\phi|<\frac{\theta}{2}\tag{31}$$

Substitution of $|\epsilon|_{\min}$ and $|\epsilon|_{\max}$ as given in equation 28 for $|\epsilon|$ in equation 31 gives the generating phasors of the boundaries of the component critical regions as

$\delta=2i \qquad 1\le i\le Q-1$

$$-\frac{1}{N}\bigg|_{\min}=\begin{cases}-\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta\theta}{2}-\phi\right)}e^{j\phi} & 0<\phi<\dfrac{\theta}{2}\\[3ex] -\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta\theta}{2}-|\phi|\right)}e^{j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}$$

$$-\frac{1}{N}\bigg|_{\max}=\begin{cases}-\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta\theta}{2}+\phi\right)}e^{j\phi} & 0<\phi<\dfrac{\theta}{2}\\[3ex] -\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta\theta}{2}+|\phi|\right)}e^{j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}\tag{32}$$

$\delta=2i+1 \qquad 0\le i\le Q-1$

$$-\frac{1}{N}\bigg|_{\min}=-\frac{\Delta}{\gamma\cos\left(\dfrac{\delta-1}{2}\theta+|\phi|\right)}e^{j\left(\phi+\frac{\theta}{2}\right)}$$

$$-\frac{1}{N}\bigg|_{\max}=-\frac{\Delta}{\gamma\cos\left(\dfrac{\delta+1}{2}\theta-|\phi|\right)}e^{j\left(\phi+\frac{\theta}{2}\right)}$$

Table I. Value of γ

T/δ	1	2	3	4	5
2	1.2732				
3	0.9549				
4	0.9003	1.2732			
5	0.7118	1.1517			
6	0.6366	1.1027	1.2732		
7	0.5386	0.9705	1.2102		
8	0.4872	0.9003	1.1763	1.2732	
9	0.4289	0.8060	1.0859	1.2349	
10	0.3935	0.7484	1.0301	1.2109	1.2732

where $\quad 0<|\phi|<\dfrac{\theta}{2}$

$\delta=T/2$

$$-\frac{1}{N}\bigg|_{\min}=\begin{cases}-\dfrac{\pi\Delta}{4\sin\phi}e^{j\phi} & 0<\phi<\dfrac{\theta}{2}\\[2ex] -\dfrac{\pi\Delta}{4\sin|\phi|}e^{j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}$$

$$-\frac{1}{N}\bigg|_{\max}=\infty$$

Similarly the describing function and generating phasor are derived for the other cases. The results are as follows. It is noticed that only for the case wherein T is an even integer $m(\tau+T/2)=-m(\tau)$ and $\Gamma(\tau+T/2)=-\Gamma(\tau)$. The absence of these qualities renders the derivation for odd values of T somewhat more involved.)

CASE 2

T$=4Q+2$ where Q is a non-negative integer. This case includes $T=2$, 6, 10. . . Equation 33 gives the expression for N

$\delta=0 \qquad N=0$

$\delta=2i \qquad 1\le i\le Q$

$$N=\frac{\gamma}{|\epsilon|}e^{-j\left(\phi+\frac{\theta}{2}\right)} \quad 0<|\phi|<\frac{\theta}{2}$$

$\delta=2i+1 \qquad 0\le i\le Q-1$

$$N=\begin{cases}\dfrac{\gamma}{|\epsilon|}e^{-j\phi} & 0<\phi<\dfrac{\theta}{2}\\[2ex] \dfrac{\gamma}{|\epsilon|}e^{-j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}\tag{33}$$

where $\gamma=4/\pi\sin\delta\theta/2$ and δ takes on the value of 0, 1, 2. ...$T/2$.

The generating phasor of the critical region is $-1/N$ and those of the boundary of critical region are

$\delta=2i \qquad 1\le i\le Q$

$$-\frac{1}{N}\bigg|_{\min}=-\frac{\Delta}{\gamma\cos\left(\dfrac{\delta-1}{2}\theta+|\phi|\right)}e^{j\left(\phi+\frac{\theta}{2}\right)}$$

$$-\frac{1}{N}\bigg|_{\max}=-\frac{\Delta}{\gamma\cos\left(\dfrac{\delta+1}{2}\theta-|\phi|\right)}e^{j\left(\phi+\frac{\theta}{2}\right)}\tag{34}$$

where $\qquad 0<|\phi|<\dfrac{\theta}{2}$

$\delta=2i+1 \qquad 0\le i\le Q-1$

$$-\frac{1}{N}\bigg|_{\min}=\begin{cases}-\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta}{2}\theta-\phi\right)}e^{j\phi} & 0<\phi<\dfrac{\theta}{2}\\[3ex] -\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta}{2}\theta-|\phi|\right)}e^{j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}$$

$$-\frac{1}{N}\bigg|_{\max}=\begin{cases}-\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta}{2}\theta+\phi\right)}e^{j\phi} & 0<\phi<\dfrac{\theta}{2}\\[3ex] -\dfrac{\Delta}{\gamma\cos\left(\dfrac{\delta}{2}\theta+|\phi|\right)}e^{j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}\tag{34}$$

$\delta=T/2$

$$-\frac{1}{N}\bigg|_{\min}=\begin{cases}-\dfrac{\pi\Delta}{4\sin\phi}e^{j\phi} & 0<\phi<\dfrac{\theta}{2}\\[2ex] -\dfrac{\pi\Delta}{4\sin|\phi|}e^{j(\phi+\theta)} & 0>\phi>-\dfrac{\theta}{2}\end{cases}$$

$$-\frac{1}{N}\bigg|_{\max}=\infty$$

CASE 3

$T=4Q+1$ where Q is a positive integer. In other words, T is 5, 9, or 13. . . The value of δ ranges from 0, 1, 2,. . . up to $(T-1)/2$

For $\delta=0 \qquad N=0$
For $\delta=1, 2, 3,. . .(T-1)/2$

$$N=\begin{cases}\dfrac{\gamma}{|\epsilon|}e^{-j\left(\frac{\theta}{2}+\phi\right)} & 0<\phi<\dfrac{\theta}{4}\\[2ex] \dfrac{\gamma}{|\epsilon|}e^{-j\phi} & \dfrac{\theta}{4}<\phi<\dfrac{\theta}{2}\\[2ex] \dfrac{\gamma}{|\epsilon|}e^{-j\left(\frac{\theta}{2}+\phi\right)} & -\dfrac{\theta}{4}<\phi<0\\[2ex] \dfrac{\gamma}{|\epsilon|}e^{-j(\phi+\theta)} & -\dfrac{\theta}{2}<\phi<-\dfrac{\theta}{4}\end{cases}\tag{35}$$

where $\gamma=(4/\pi)\cos\theta/4\sin\delta\theta/2$.

Its negative reciprocal, $-1/N$, gives the expression for the generating phasor, and the generating phasor of the boundary of the critical region is

For $\delta = 1, 2, 3, \ldots T-1/2$

$$-\frac{1}{N}\bigg|_{\min} = \begin{cases} -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+\phi\right)} e^{j\left(\frac{\theta}{2}+\phi\right)} & 0<\phi<\dfrac{\theta}{4} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}\theta-\phi\right)} e^{j\phi} & \dfrac{\theta}{4}<\phi<\dfrac{\theta}{2} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+|\phi|\right)} e^{j\left(\frac{\theta}{2}+\phi\right)} & -\dfrac{\theta}{4}<\phi<0 \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}\theta-|\phi|\right)} e^{j(\theta+\phi)} & -\dfrac{\theta}{2}<\phi<-\dfrac{\theta}{4} \end{cases} \tag{36}$$

$$-\frac{1}{N}\bigg|_{\max} = \begin{cases} -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}\theta-\phi\right)} e^{j\left(\frac{\theta}{2}+\phi\right)}, & 0<\phi<\dfrac{\theta}{4} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+\phi\right)} e^{j\phi}, & \dfrac{\theta}{4}<\phi<\dfrac{\theta}{2} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}\theta-|\phi|\right)} e^{j\left(\frac{\theta}{2}+\phi\right)}, & -\dfrac{\theta}{4}<\phi<0 \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+|\phi|\right)} e^{j(\theta+\phi)}, & -\dfrac{\theta}{2}<\phi<-\dfrac{\theta}{4} \end{cases}$$

CASE 4

$T = 4Q+3$ where Q is a non-negative integer. (T is 3, 7, or 11...) In this case δ takes on the values of 0, 1, 2,... up to $T-1/2$.

For

$\delta = 0$; $N = 0$

For

$\delta = 1, 2, 3, \ldots T-1/2$

$$N = \begin{cases} \dfrac{\gamma}{|\epsilon|} e^{-j\left(\phi+\frac{\theta}{2}\right)} & 0<\phi<\dfrac{\theta}{4} \\ \dfrac{\gamma}{|\epsilon|} e^{-j\phi} & \dfrac{\theta}{4}<\phi<\dfrac{\theta}{2} \\ \dfrac{\gamma}{|\epsilon|} e^{-j\left(\phi+\frac{\theta}{2}\right)} & -\dfrac{\theta}{4}<\phi<0 \\ \dfrac{\gamma}{|\epsilon|} e^{-j(\phi+\theta)} & -\dfrac{\theta}{2}<\phi<-\dfrac{\theta}{4} \end{cases} \tag{37}$$

where $\gamma = 4/\pi \cos\theta/4 \sin\delta\theta/2$.

The expression of the generating phasor is given as the negative reciprocal of equation 37, namely $-1/N$. That of the boundary of the critical region is

$\delta = 1, 2, 3, \ldots (T-1)/2$

$$-\frac{1}{N}\bigg|_{\min} = \begin{cases} -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+\phi\right)} e^{j\left(\frac{\theta}{2}+\phi\right)} & 0<\phi<\dfrac{\theta}{4} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}-\phi\right)} e^{j\phi} & \dfrac{\theta}{4}<\phi<\dfrac{\theta}{2} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+|\phi|\right)} e^{j\left(\frac{\theta}{2}+\phi\right)} & -\dfrac{\theta}{4}<\phi<0 \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}\theta-|\phi|\right)} e^{j(\theta+\phi)} & -\dfrac{\theta}{2}<\phi<-\dfrac{\theta}{4} \end{cases}$$

$$-\frac{1}{N}\bigg|_{\max} = \begin{cases} -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}\theta-\phi\right)} e^{j\left(\frac{\theta}{2}+\phi\right)} & 0<\phi<\dfrac{\theta}{4} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+\phi\right)} e^{j\phi} & \dfrac{\theta}{4}<\phi<\dfrac{\theta}{2} \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta+1}{4}\theta-|\phi|\right)} e^{j\left(\frac{\theta}{2}+\phi\right)} & -\dfrac{\theta}{4}<\phi<0 \\ -\dfrac{\Delta}{\gamma\cos\left(\dfrac{2\delta-1}{4}\theta+|\phi|\right)} e^{j(\theta+\phi)} & -\dfrac{\theta}{2}<\phi<-\dfrac{\theta}{4} \end{cases} \tag{38}$$

As $|\epsilon|$ varies through its limiting values and ϕ varies from $-\theta/2$ to $\theta/2$, $-1/N$ generates a critical region on the complex plane for each pair of δ and T. This region is named the component critical region of that pair of δ and T. It can be best constructed by using $-1/N|_{\min}$ and $-1/N|_{\max}$. The critical region for a given period T is the union of the component critical region of that T and for all possible δ's.

Examination of the expression for the describing function N and the conditions on ϵ to give each value of δ leads to the following conclusions:

1. The describing function of the nonlinear element under investigation is a function of frequency, amplitude, and the phase of the input sinusoid ϵ, and inactive zone Δ.

2. The nonlinear element is essentially a phase lag device, the amount of phase lag depending solely upon T and ϕ. The range of phase lag is from 0 to $\theta = 2\pi/T$ for T of even integer value and from $\theta/4$ to $3\theta/4$ for T of odd integer value.

3. The magnitude of N is a function of $|\epsilon|$, T, ϕ, and Δ. For given values of T and ϕ, δ is a discontinuous jump function of $|\epsilon|/\Delta$. The discontinuities occur at the $|\epsilon|_{\min}/\Delta$'s and $|\epsilon|_{\max}/\Delta$'s at which the value of δ jumps from one integer to the next. In consequence, the magnitude of N is also a discontinuous function of $|\epsilon|/\Delta$, having the same discontinuities as δ. Between any two consecutive discontinuities, the magnitude of describing function N is inversely proportional to $|\epsilon|$. Since N is zero for $\delta = 0$, it has a sharp amplitude cutoff.

For illustration, Fig. 4 shows the magnitude of N as a function of $|\epsilon|/\Delta$, given $T = 8$ and $\phi = \pi/18$.

A study of the generating phasors reveals the following facts:

1. The generating phasors for ϕ and $-\phi$ form a pair of mirror images about the $-(\pi-\theta/2)$ radial line of the complex plane, hence the critical region is symmetrical about the $-(\pi-\theta/2)$ radial line.

2. From the range of phase lag of N, it is seen that the critical region of even (odd)

Fig. 9. The critical region of T=2

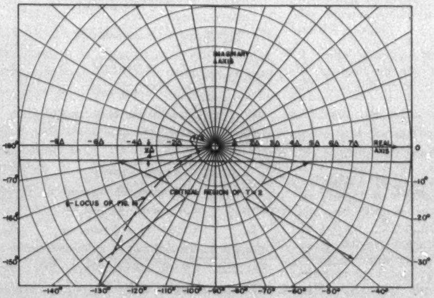
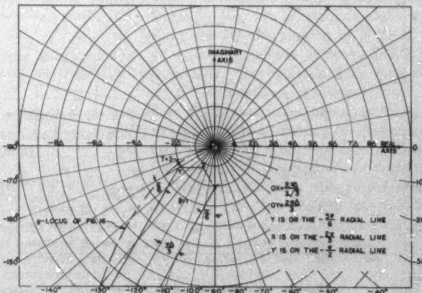


Fig. 10. The critical region of T=3



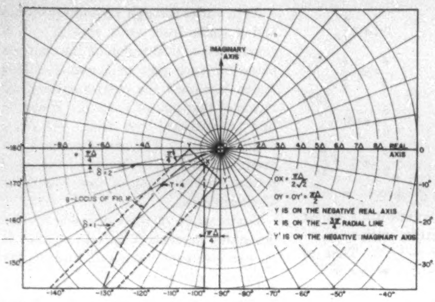


Fig. 11. The critical region of T=4

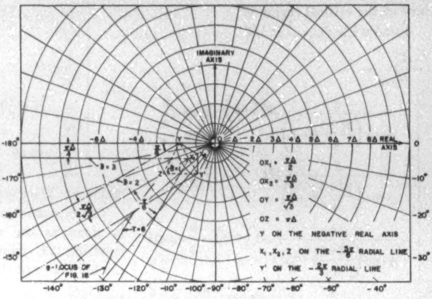


Fig. 13. The critical region of T=6

values of T lies on the sector bounded by $-\pi[-(\pi-\theta/4)]$ and $-(\pi-\theta)[-(\pi-3\theta/4)]$ radial lines. It is in the third quadrant except the case of $T=2$, wherein its critical region lies in the lower half-plane.

3. Each elementary phasor constituting the generating phasor of boundary generates a straight line on the complex plane, hence the boundary of critical region consists of segments of straight lines.

4. The length of generating phasors of boundary is directly proportional to the value of the contactor inactive zone Δ. Hence, as the value of Δ increases, the critical region moves away from the origin of the plane. This property, as will be seen, is very useful in providing system stability.

5. For certain values of $|\epsilon|$ and ϕ, $-1/N$ is infinite, hence the critical region extends to infinity. Near the origin of the complex plane the critical region is bounded by origin-end boundary to be determined in the next section.

6. As the period T increases through even (or odd) values, the critical regions gets narrower, and finally approaches the $-\pi$ radial line, that is, the negative real axis of the complex plane as T approaches infinity.

7. For each component critical region there are three corresponding characteristic values, namely, T, δ, and γ. The value of γ is a function of T and δ only. Explicitly γ is

For T of even integer value

$$\gamma=4/\pi \sin \delta\theta/2 \tag{39}$$

For T of odd integer value

$$\gamma=4/\pi \cos \theta/4 \sin \delta\theta/2 \tag{40}$$

where $\theta=2\pi/T$.

The value of γ for $2 \leq T \leq 10$ is show in Table I.

Origin End Boundary

As has been shown, $-1/N|_{\min}$ for a given T is a function of δ; the origin end boundary of the critical region of a given T is the locus generated by $\overset{\min}{\delta}[-1/N|_{\min}]$; the minimum of $-1/N|_{\min}$ is in the sense of minimum length of $-1/N|_{\min}$ over all possible δ's.

CASE 1

$T=4Q$

The expression of $-1/N|_{\min}$ is rewritten from equation 32 by a proper substitution as follows

$$-\frac{1}{N}\bigg|_{\min}=\frac{\pi\Delta}{4 \sin\dfrac{\delta\theta}{2} \cos\left(\dfrac{\delta\theta}{2}-\psi\right)}e^{j(\psi-\pi)} \tag{41}$$

where

$$\psi=\begin{cases}\phi & 0<\phi<\dfrac{\theta}{2} \text{ for even } \delta\\ \phi+\dfrac{\theta}{2} & 0>\phi>-\dfrac{\theta}{2} \text{ for odd } \delta\end{cases}$$

Although the value of ψ ranges from 0 to θ, because of the symmetry of the critical region about the $-(\pi-\theta/2)$ radial line, only $0<\psi<\theta/2$ needs to be considered. Equation 41 only defines the phasor for this range $(0, \theta/2)$.

The length of $-1/N|_{\min}$ is given then

$$\left|-\frac{1}{N}\bigg|_{\min}\right|=\frac{\pi\Delta}{4 \sin\dfrac{\delta\theta}{2} \cos\left(\dfrac{\delta\theta}{2}-\psi\right)} \tag{42}$$

Minimizing this length with respect to δ is equivalent to maximizing $\sin \delta\theta/2 \cos (\delta\theta/2-\psi)$, with respect to δ, for all ψ, $0<\psi<\theta/2$. Since $\sin \psi$ is independent of δ, then

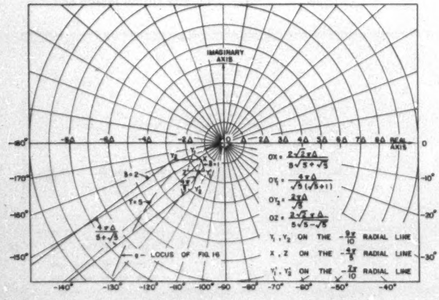


Fig. 12. The critical region of T=5

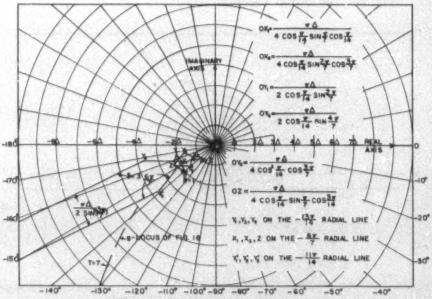


Fig. 14. The critical region of T=7

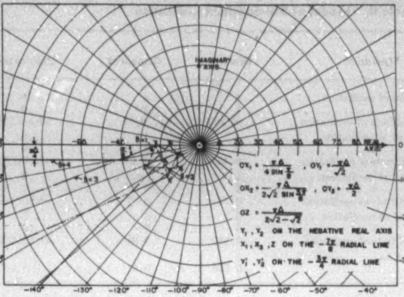

IMAGINARY AXIS
REAL AXIS
−180°
−170°
−160°
−150°
0
−10°
−20°
−30°
−140°
−130°
−120°
−110°
−100°
−90°
−80°
−70°
−60°
−50°
−40°
−8Δ
−4Δ
δ=1
δ=2
δ=3
δ=4
$\frac{\pi\Delta}{4}$
2Δ
3Δ
4Δ
5Δ
6Δ
7Δ
8Δ
$OX_1 = \frac{\pi\Delta}{4 \sin\frac{\pi}{8}}$, $OY_1 = \frac{\pi\Delta}{\sqrt{2}}$
$OX_2 = \frac{\pi\Delta}{2\sqrt{2}\sin\frac{3\pi}{8}}$, $OY_2 = \frac{\pi\Delta}{2}$
$OZ = \frac{\pi\Delta}{2\sqrt{2-\sqrt{2}}}$
Y_1, Y_2 ON THE NEGATIVE REAL AXIS
X_1, X_2, Z ON THE $-\frac{7\pi}{8}$ RADIAL LINE
Y_1', Y_2' ON THE $-\frac{3\pi}{4}$ RADIAL LINE

Fig. 15 (above). The critical region of T=8

$$\overset{\min}{\delta}\left|\frac{1}{N}\right|_{\min} \rightleftharpoons \overset{\max}{\delta}\left[\sin\frac{\delta\theta}{2}\cos\left(\frac{\delta\theta}{2}-\psi\right)\right]$$

$$= \overset{\max}{\delta}\frac{1}{2}[\sin(\delta\theta-\psi)+\sin\psi]$$

$$\rightleftharpoons \overset{\max}{\delta}\sin(\delta\theta-\psi) \quad (43)$$

Since $\theta=2\pi/T$, $0<\psi<\pi/T$, and δ can only take integer values of 0, 1, 2,. . .$T/2$, then to maximize $\sin(\delta\theta-\psi)$ with respect to δ is equivalent to making $\delta\theta-\psi$ closest to $\pi/2$. This is achieved for all ψ, $0<\psi<\pi/T$ by putting $\delta=T/4=Q$.

Similarly it can be proved that the origin end boundaries of the critical regions of $T=4Q+2$, $T=4Q+1$, and $T=4Q+3$ are respectively generated by

$$-1/N|_{\min} \text{ of } \delta=T+2/4$$
$$=Q+1,\ \delta=T-1/4=Q$$

$$\delta=T+1/4=Q+1$$

They are shown in Figs. 5 through 8.

Critical Regions

By use of the generating phasors the critical regions of $2\leq T\leq 8$ are constructed in Figs. 9 through 15 for illustrative purposes, with essential dimensions shown on the figures. For convenience they are plotted to a Δ-scale.

System Stability

The stability of the nonlinear system is briefly defined in the following way: If the system is capable of maintaining a self-sustained oscillation then the system is unstable, otherwise it is stable.

Treating N as a linear transfer function, the sinusoidal response characteristic equation of the over-all servo system will be

$$\frac{C(ju)}{R(ju)}=\frac{Ng(ju)}{1+Ng(ju)} \quad (44)$$

Consideration of the nature of N and $g(ju)$ and equation 44 shows that self-sustained periodic oscillations may exist when

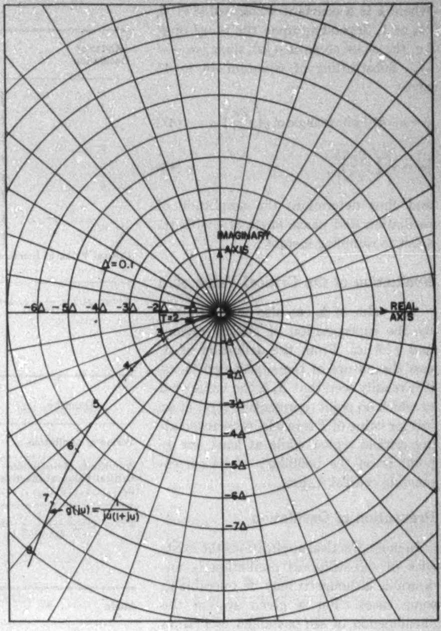


Fig. 16 (right). The g-locus of the example

$$1+Ng(ju)=0 \quad (45)$$

or $g(ju)=-1/N$ and the system is unstable.

The solution of equation 45 can be best obtained by employing the Nyquist g-plot as it is extensively used in treating linear servomechanisms. Equation 45 reveals that the critical point on the complex plane is $-1/N$ instead of $-1+j0$ point as in the case of linear servo systems.

The procedure of determining the system stability consists of:

1. Plotting $g(ju)$ on the complex plane for the whole frequency range $0\leq u\leq\infty$. Integer value of $T=2\gamma/u$ is used as a parameter on the g-locus.

2. Superposing on the same plane a family of critical regions, the critical region of $T=2, 3, 4, 5\ldots$

From the relative position of the g-locus and the critical regions the following conclusions are reached:

1. Should the point representing $g(j2\pi/T)$ for some T on the complex plane fall within the critical region of this T (of course T is an integer ≥ 2), it means that there exists T, δ, and $|\epsilon|$, such that $g(j2\gamma/T)=-1/N|_{T,|\epsilon|,\delta}$. In consequence, self-sustained oscillation at T and $|\epsilon|$, characterized by δ, may exist and the system is unstable. Should $g(j2\gamma/T)$ for more than one value of T fall within their respective critical regions, equation 45 has more than one solution, i.e., the system has more than one mode of oscillation.

2. Otherwise the system is stable.

Determination of Modes of Self-Sustained Oscillation

Should the superposition show that for some integer value of T, say T', $g(j2\pi/T')$ is on the critical region of T', then the self-sustained oscillation of period T' may exist. Since the critical region for T' is the union of its component regions of $\delta=1, 2, 3,\ldots T'/2$ or $T'-1/2$, depending on whether T' is even or odd, one is able to determine by the position of $g(j2\pi/T')$ on what component region $g(j2\pi/T')$ is located. This determines the value of δ. (It frequently occurs that $g(j2\pi/T')$ lies on the intersection, namely the overlapping part of two or more component regions. If this is the case, then each of these values of δ represents one mode of oscillation.) Knowing the values of δ and $T=T'$, one can find the corresponding expression for $-1/N$ and the value of γ. $-1/N$ has the following form

$$-\frac{1}{N}=-\frac{|\epsilon|}{\gamma}e^{j(\phi+k\theta)} \quad (46)$$

where k is a constant whose value is 0, $\frac{1}{2}$ or 1, depending upon the range of ϕ, i.e., the phase angle of $g(ju)$, since $g=-1/N$. Substituting of equation 46 in 45 gives

$$\phi=\pi-k\theta+\text{phase angle of } g\left(j\frac{2\pi}{T'}\right) \qquad (47)$$

$$|\epsilon|=\gamma\left|g\left(j\frac{2\pi}{T'}\right)\right| \qquad (48)$$

and thus determines the amplitude of oscillation and phase angle of ϵ with respect to sampling instant.

Evaluation of the Critical Value of Δ

The value of Δ_c can also be determined by the superposition of critical regions and g-locus. From the definition of Δ_c and the nature of the superposed locus, it is readily seen that Δ_c is the value of Δ at which no point representing $g(ju)$ of an integer value of T is inside its corresponding critical region while at least one g-point is on the boundary of its corresponding critical region.

Prevention of Oscillation

In most practical applications the existence of self-sustained oscillation is undesirable although it may be tolerable in some cases. For a given system the phenomenon of self-sustained oscillation can be eliminated by employing either or both of the following methods:

1. By the use of larger value of the inactive zone of the contractor Δ ($>\Delta_c$).
2. By the use of a proper compensating network.

The underlying principle of preventing oscillation is to adjust the relative positions of the g-locus and the critical regions on the complex plane, such that no point of $g(ju)$ of integer period is within the corresponding critical region.

Since the distance of the origin end boundary of the critical region from the origin of the complex plane is directly proportional to the value of Δ, the use of a larger Δ simply increases this distance, and moves the critical region away from the origin, and thus avoids the proper intersection of these two configurations.

It is recalled that the relation between the dimensionless inactive zone and the actual dimensional quantities is $\Delta=(\theta_\Delta)/(KT_s^\lambda)$. The increase in the value of Δ can be achieved by one or by the combined effects of the following:

1. The increase of the actual inactive zone of the contactor θ_Δ.
2. The decrease of the over-all forward loop gain K.
3. The decrease of sampling period T_s, i.e., the increase in the sampling frequency $2\gamma/T_s$.

Table II. Modes of Self-Sustained Oscillation

Modes of Oscillation	Quantity	Predicted	Computed	Discrepancy
1	Period of oscillation	4	4	0
	δ	1	1	0
	Amplitude of oscillation	0.308	0.31	within 1%
2	Period of oscillation	4	4	0
	δ	2	2	0
	Amplitude of oscillation	0.435	0.432	within 1%
3	Period of oscillation	5	5	0
	δ	2	2	0
	Amplitude of oscillation	0.571	0.57	within 1%
4	Period of oscillation	6	6	0
	δ	3	3	0
	Amplitude of oscillation	0.840	0.84	within 1%
Critical Value of Inactive Zone Δ_c				
Δ_c		0.526	0.51	+3.14%

Note: $g(s)=\frac{1}{s(1+s)}$, $\Delta=0.1$

Table III. Mode of Oscillation

Quantity	Predicted	Computed	Discrepancy per cent
Period of oscillation	4	4	0
δ	1	1	0
Amplitude of oscillation	0.5468	0.50	+9.36
Critical Value of Inactive Zone			
Δ_c	0.4848	0.476	1.85

Note: $g(s)=\frac{1}{s(1+0.2s)}$ $\Delta=0.4$

Table IV. Modes of Self-Sustained Oscillation

Mode	Quantity	Predicted	Computed	Discrepancy
Case 1	$\zeta=0.5$ $u_n=10$			
1	Period of oscillation	2	2	0
	δ	1	1	0
	Amplitude of oscillation	0.4246	0.46	−7.7%
2	Period of oscillation	4	4	0
	δ	2	2	0
	Amplitude of oscillation	0.8205	0.94	−12.7%
Case 2	$\zeta=1.0$ $u_n=10$			
1	Period of oscillation	2	2	0
	δ	1	1	0
	Amplitude of oscillation	0.3689	0.375	− 1.63%
2	Period of oscillation	3	3	0
	δ	1	1	0
	Amplitude of oscillation	0.4361	0.444	−1.78%
3	Period of oscillation	4	4	0
	δ	2	2	0
	Amplitude of oscillation	0.7910	0.875	−9.60%
Critical Value of Inactive Zone, Δ_c				
Case 1	$\zeta=0.5$	0.4703	0.475	within 1%
Case 2	$\zeta=1.0$	0.4977	0.495	within 1%

Note: $g(s)=:\frac{1}{s\left(\frac{s^2}{u_n^2}+\frac{2\zeta s}{u_n}+1\right)}$ where $u_n=10$; and $\Delta=0.1$

The employment of a compensating network will modify the shape of the g-locus on the complex plane. A proper selection of a compensating network will swing the g-locus away from the critical regions and thus reduce the critical value of Δ.

Numerical Example

For illustration, a simple example is given. Let $g(s)=1/s(1+s)$ and $\Delta=0.1$, dimensionless.

The g-locus is constructed in Fig. 16. Superposition of the critical regions reveals that $g(ju)$ of $T=4$, 5 and 6 are in their respective critical regions while all other g-points are outside their corresponding regions. (This is not shown because of the complicated diagram which would result if all critical regions were included. Instead, the g-locus is superposed on Figs. 9 through 14. To facilitate the superposition of critical regions on the g-locus, transparent paper is used in actual application.) Thus the system

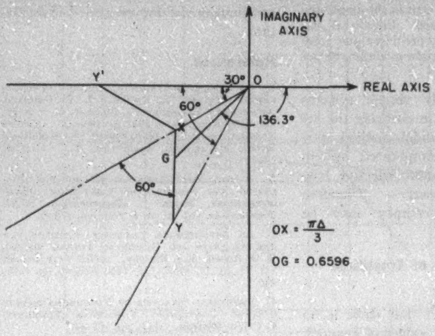


Fig. 17. Determination of Δ_c

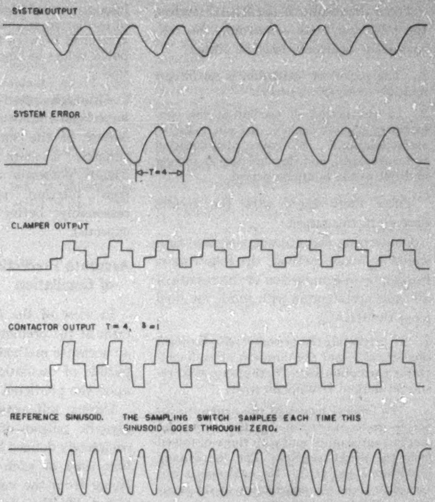


Fig. 18 (right). REAC computer graph showing a mode of self-sustained oscillation of the system with transfer function $g(s)=1/s(1+s)$ and $\Delta=0.1$

is unstable and may oscillate at a period of 4, 5, or 6 dimensionless time units.

Further examination of the superposed loci reveals that the g-point of $T=6$ is in the component region of $\delta=3$, the g-point of $T=5$ is in the component region for $\delta=2$, and the g-point of $T=4$ is in the intersection of the component regions of $\delta=1$ and $\delta=2$. Consequently the system has four possible modes of oscillation.

For T=6 and $\delta=3$, γ is $4/\pi$ and, since the phase angle of $g(ju)$ of $T=6$, is -136.3 degrees, which is in the range of -120 to -150 degrees, $-1/N$ is given as

$$-\frac{1}{N}=-\frac{|\epsilon|}{\gamma}e^{j\left(\frac{\pi}{3}+\phi\right)}$$

Equating the last expression with $g(ju)$ of $T=6$, which is $0.6596e^{-j136.3°}$ gives

$$\phi=-16.3 \text{ degrees}$$

$$|\epsilon|=0.840$$

Similarly it is evaluated that the amplitudes of self-sustained oscillation at $T=5$, $\delta=2$; $T=4$, $\delta=1$; and $T=4$, $\delta=2$ are respectively 0.571, 0.308, and 0.435 dimensionless.

Since the phase lag of $g(ju)$ of $T\geqslant 7$ is less than 132° while the critical regions for $T\geqslant 7$ are within the sector bounded by $-135°$ and $-180°$ radial lines, g-points of $T\geqslant 7$ will not fall inside their respective critical regions no matter what the value of Δ is. An inspection of the superposed loci shows that Δ_c is that value of Δ at which $g(ju)$ of $T=6$ is on the origin end boundary of the critical region while all other g-points of integer T are outside their corresponding regions. The value of Δ_c can be read graphically, or more precisely, it can be calculated from the known value of $g(ju)$ and the geometric relation as given in Fig. 17 in which YXY' is the origin end boundary of the critical region of $T=6$ and OG is the phasor representing $g(ju)$ of $T=6$. At Δ_c, G is on the boundary YXY' and the calculated value of Δ_c is 0.526.

Accuracy of the Method

The accuracy of the sinusoidal response method depends primarily upon the type of nonlinearity and the low-pass filtering property of the linear part of the servo system. No general statement concerning the accuracy can be made. The accuracy of the method as applied to this particular type of servomechanism is obtained by actual comparison of the predicted results with those obtained from a Reeves Electronic Analogue Computer (REAC).

The comparison of results for the preceding example is shown in Table II. The discrepancy is rather small, mainly because of the effectiveness of the low-pass characteristic of $g(s)=1/[s(1+s)]$. One of the typical graphs obtained from the REAC is shown in Fig. 18.

Table III gives the comparison of results for a system having

$$g(s)=\frac{1}{s(1+0.2s)}$$

$$\Delta=0.4$$

Table IV gives the comparison of results for a system having

$$g(s)=\frac{1}{s\left(\frac{s^2}{u_n^2}+\frac{2\zeta s}{u_n}+1\right)}$$

of resonant type and $\Delta=0.1$.

Table V gives that for a system having

$$g(s)=\frac{1}{s(1+s)(1+0.5)} \text{ and } \Delta=0.2.$$

Table V. Modes of Self-Sustained Oscillation

Mode	Quantity	Predicted	Computed	Discrepancy
1	Period of oscillation	6	6	0
	δ	2	2	0
	Amplitude of oscillation	0.6443	0.64	within 1%
2	Period of oscillation	7	7	0
	δ	3	3	0
	Amplitude of oscillation	0.9154	0.90	+1.7
3	Period of oscillation	8	8	0
	δ	3	3	0
	Amplitude of oscillation	1.0964	1.10	within 1%
4	Period of oscillation	8	8	0
	δ	4	4	0
	Amplitude of oscillation	1.1867	1.2	−1.1%
Critical Value of Inactive Zone				
Δ_c		0.7233	0.73	within 1%

Note: $g(s)=\frac{1}{s(1+s)(1+0.5s)}$, $\Delta=0.2$

From the results of the REAC studies, the following facts concerning the self-sustained oscillation may be stated:

1. The period of oscillation is an integer multiple of sampling period.
2. In one period of oscillation the contactor exerts one positive and one negative correction signal. Moreover, their duration times are equal, and they are also an integer multiple of the sampling period.

These facts agree with the results derived in this paper.

Concerning the accuracy of the sinusoidal response method, the following is learned from comparison of the results of all cases investigated with those obtained from the REAC:

1. In predicting the period of self-sustained oscillation T and the duration of each contactor correction signal δ, the sinusoidal response method exhibits no error.
2. The critical values of the contactor inactive zone predicted by the sinusoidal method agree quite well with those obtained from the REAC studies. The discrepancy is within 5 per cent. No evidence is observed to show the dependence of the accuracy on the type of system.
3. The error in the prediction of amplitudes of oscillation by the sinusoidal method depends primarily upon the effectiveness of the low-pass filtering action of the servo system. The more effective it is, the better the accuracy will be. In some cases the error is within 1 per cent, but in the case wherein the servo system has a resonant type of transfer function in the forward loop, an error of 13 per cent in the amplitude of oscillation is observed. In general, the agreement between the predicted and computed values of amplitude is within 10 per cent.

Although relatively simple systems have been employed in checking the accuracy of the sinusoidal method, fortunately a more complicated system usually possesses a more effective low-pass filtering property. Therefore reasonably better accuracy may be expected.

Accurate Prediction of Amplitude of Oscillation

In view of the fact that there is no error in the predicted values of T and δ, an accurate method of predicting the amplitude of oscillation is achieved, based upon the predicted values of T and δ.

During self-sustained oscillation, the periodic output of the contactor Γ is completely defined by period T and duration time of each correction signal δ. Hence from the values of T and δ predicted by the sinusoidal method, the form of Γ is obtained. In consequence, the linear differential equation

$$\frac{1}{g(D)}c(\tau)=\Gamma(\tau)$$

can be solved for the exact amplitude of oscillation by conventional methods and matching the end conditions at the discontinuities of Γ for one period of oscillation.

References

1. ON SOME NON-LINEAR PHENOMENA IN REGULATORY SYSTEMS (in Russian), **L. C. Goldfarb.** *Avtomatika i Telemechanika*, Moscow, U. S. S. R., vol. 8, no. 5, Sept.-Oct. 1947, pp. 349–83. (English Translation: *Report 1691*, United States National Bureau of Standards, Washington, D. C., May 29, 1952.)

2. A FREQUENCY RESPONSE METHOD FOR ANALYZING AND SYNTHESIZING CONTACTOR SERVOMECHANISMS, **Ralph J. Kochenburger.** *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

3. AN APPROXIMATE TRANSFER FUNCTION FOR THE ANALYSIS AND DESIGN OF PULSED SERVOS, **R. G. Brown, G. J. Murphy.** *AIEE Transactions*, vol. 71, pt. II, 1952 (Jan. 1953 section), pp. 435–440.

4. SINUSOIDAL ANALYSIS OF FEEDBACK-CONTROL SYSTEMS CONTAINING NONLINEAR ELEMENTS, **E. Calvin Johnson.** *Ibid.*, pp. 169–81.

5. COULOMB FRICTION IN FEEDBACK CONTROL SYSTEMS, **Vinton B. Haas, Jr.** *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 119–23.

6. LIMITING IN FEEDBACK CONTROL SSYTEMS, **Ralph J. Kochenburger.** *AIEE Transactions*, vol. 72, pt. II, July 1953, pp. 180–94.

7. THE EFFECTS OF BACKLASH AND OF SPEED-DEPENDENT FRICTION ON THE STABILITY OF CLOSED-CYCLE CONTROL SYSTEMS, **A. Tustin.** *Journal*, Institution of Electrical Engineers, London, England, vol. 94, pt. IIA, 1947, pp. 143–51.

8. CONTACTOR SERVOMECHANISMS EMPLOYING SAMPLED-DATA, **C. K. Chow.** *Thesis*, Cornell University, Ithaca, N. Y., 1953.

9. AN ANALYSIS OF RELAY SERVOMECHANISMS, **David A. Kahn.** *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 1079–88.

10. INTRODUCTION TO NONLINEAR MECHANICS (book), **N. Kryloff, N. Bogoliuboff.** Princeton University Press, Princeton, N. J., 1943.

11. INTRODUCTION TO NON-LINEAR MECHANICS, (book), **N. Minorsky.** J. W. Edwards Bros., Ann Arbor, Mich., 1947.

◆

Discussion

Frederick A. Russell (Newark College of Engineering, Newark, N. J.): That Dr. Chow and the writer were simultaneously engaged in solving the same problem is an interesting coincidence.[1] It is not surprising, perhaps, when one considers the present interest in radio pulse-controlled servo systems and the attractive simplicity and lightness of on-off control.[2] Although the chopper-bar galvanometer relay has been used for some time in systems of this general nature for relatively slow process-control work,[3] Dr. Chow is the first to publish a method which can be used for the design of stable systems. The use of transfer functions and application of the Barkhausen criterion for oscillation are devices which should appeal to engineers used to working with linear systems.

MAGNITUDE-FREQUENCY AND PHASE-FREQUENCY CURVES

The writer found that presentation of the Barkhausen equation given as equation 45

$$1+N(j\omega;\theta_e;\phi)KG(j\omega)=0 \tag{49}$$

or

$$-\frac{1}{N(j\omega;\theta_e;\phi)}=KG(j\omega) \tag{50}$$

is greatly simplified if put in logarithmic form, and reals and imaginaries are equated. Thus, dropping the functional notation

$$\frac{1}{|N|}=K|G| \tag{51}$$

$$-(\pi-\phi_N)=\phi_G \tag{52}$$

On the right are, respectively, the magnitude $K|G|$ and the phase shift ϕ_G for the linear portion of the system (which is permitted to have additional synchronized linear samplers and clampers[4]). On the left are the reciprocal of the magnitude $1/|N|$ and $-\pi$ plus the phase shift ϕ_N of the describing function. Here ϕ_G and ϕ_N are in addition to the symbols used by Chow.

These equations are simply presented as magnitude versus frequency and phase-shift versus frequency curves. Logarithmic scales can be used if desired, in which case the linear system curves are the familiar Bode diagram, and design alterations in $KG(p)$ can readily be made thereon. For brevity, only nonlogarithmic scales will be discussed here.

Consider first the phase graph. For each even-integer T value, $-(\pi-\phi_N)$ has a range from $-\pi$ to $-(\pi-\theta)$. The phase graph for all even-integer T will therefore be a series of vertical lines, the upper and lower bounds of which are indicated by radial lines in Figs. 5 and 6. For odd-integer T there will be another series of vertical lines within bounds shown in Fig. 7. A few of these are shown in Fig. 19 together with the straight-line boundaries and the boundaries' bisector, $-(\pi-\theta/2)$.

The Barkhausen criterion of equations 51 and 52 cannot be satisfied unless there is an intersection of the ϕ_G curve and a ϕ_N line, and then only if the magnitude relationship is satisfied also at the same value of signal angular frequency ω. Hence there can be no sustained oscillation except within the band of frequencies wherein ϕ_G lies within the "potentially unstable phase region" between $-\pi$ and $-(\pi-\theta)$.

Consider now the magnitude graph. The range of values for $1/|N|$ at a given T depends upon, first, the type of number T is, for example, even or odd integer; and, second, the depth of penetration of ϕ_G into the ϕ_N line at that T value. In Fig. 19, at $T=22$, the ϕ_G line barely touches the ϕ_N line; hence the ranges are shrunk to mere points, each of which represents a possible

Table VI. Absolute Minima for T

T	$1/(\|N\|\theta_\Delta)$
4	1.110
8	1.202
12	1.283
16	1.340
20	1.374
24	1.403
28	1.421
32	1.440
36	1.450
40	1.462

As $T\rightarrow\infty$, $1/|N\rightarrow\infty\theta_\Delta\pi/2$

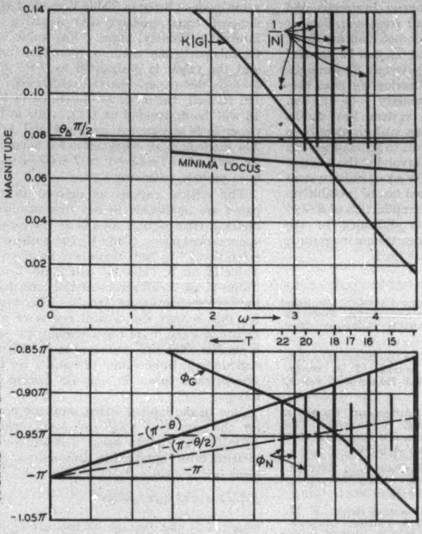


Fig. 19. Illustrative amplitude and phase versus frequency graphs for integer values of T for sampled-data on-off servo. Graphs drawn for T=0.1 second, θ_Δ= 0.05 radian, F(jω) arbitrary. System is unstable at ω=3.14 or T=20

mode of operation. Magnitude ranges exist only where there is a phase intersection. If the ϕ_G line were shifted downward, these points would become vertical lines, and there would be a possibility of a $K|G|$ curve intersection. Simultaneous intersections on magnitude and phase graphs would indicate instability for that particular mode of operation δ.

The $1/|N|$ lines increase in length as ϕ_G moves downward, eventually overlapping, so that multiple modes of unstable operation at a single value of frequency are possible, as pointed out by Dr. Chow. The lines reach their maximum length, whose lower ends are also the minimum values, when ϕ_G crosses the line $-(\pi-\theta/2)$, as at $T=18$, and then decrease in length as ϕ_G proceeds further downward. In Fig. 19 there are no possible modes of oscillation for T of 21 and 15 because these ϕ_N lines are not intersected by ϕ_G.

The $1/|N|$ ranges, as stated, also depend upon the kind of number T is, and are longest and reach lower minima for integer T divisible by 4. Thus the absolute minima will be for such T when $\phi_G=-(\pi-\theta/2)$, and these minima are given in Table VI.

These minima always occur for the middle mode, that is, the one with equal on and off intervals. Hence, as the gain is increased and $K|G|$ moved upward, instability is likely to occur first in one of these modes. The locus given in Table VI constitutes the boundary of the "potentially unstable magnitude region," for if the $K|G|$ curve lies beneath it, no instability can occur. It should be noted that this locus can be approximated by the straight line at $\theta_\Delta\pi/2$.

We are now prepared to state a stability criterion for the design of systems having T_s appreciably smaller than the largest time constant in $G(p)$: If the curves of magnitude and phase for the linear portion of the system are not, over any band of frequencies, simultaneously within their respective potentially unstable regions, the system is stable. It should be pointed out that this criterion requires drawing only the lines at $-\pi$ and $-(\pi-\theta)$ on the phase graph, and the locus tabulated on the magnitude graph. For most purposes, where considerable gain and phase margins are desired, the straight line at $-\theta_\Delta/2$ may be considered the magnitude region boundary. Margins are laid out merely by drawing lines parallel to the boundaries, and outside of the potentially unstable regions.

MODES OF OSCILLATION

I disagree with Dr. Chow in two respects which greatly simplified his analysis. He assumed that the sampler, clamper, and contactor output are always periodic at signal frequency; and he stated that one or more integrations in $G(p)$ indicated that positive and negative contactor ON intervals must be of equal duration. In the first place, equal ON intervals do not assure symmetrical error signals. For example, the system $G(p)=1/p(1+p)$, with $T_s=0.5$ and $\theta_\Delta=0.1$, when subjected to an initial error signal of 0.2 has a transient response which converges rapidly to a steady-state mode of operation consisting of three positive ON intervals, zero OFF intervals, three negative ON intervals, and 1 OFF interval. The mode of operation may therefore be termed "mode 3031" (the number of posituve ON intervals arbitrarily being given first), which indicates a period $T=7$ sampling cycles. A graph of the error signal for this case shows that it possesses an average component of about 0.052 or one-half the dead zone. This offset occurs because the contactor can close only at sampling instants, and none occurs with a separation of π radians from another for a mode of this sort. One might say that the eighth pulse, delayed by one OFF period, creates the equivalent of a system disturbance, thereby causing an offset error much like an average torque applied to the output member.

Unequal ON intervals can be unequal by only one period T_s, and will cause the same sort of effect. Admission of these cases doubles the number of possible modes for $T=4Q+1$ and $4Q+3$, and makes the mode number designation δ inadequate. There are then twice as many possible ranges for $1/|N|$ for odd-integer T. The phase boundaries are the same, and the minima will all lie within the line $\theta_\Delta\pi/2$.

When the ratio of sampling frequency to signal frequency is not an integer, the sampler becomes an "almost periodic function" of signal frequency,[8] and likewise the clamper and contactor outputs. They will then be periodic with period CT where the integer C is 1 if the frequency ratio is an integer; 2 if an odd half-integer; 3 if a third of an integer; 4 if an odd quarter of an integer, etc. C approaches infinity as the frequency ratio approaches an irrational number. Consider, for example, the frequency ratio 9.5. Possible modes of operation are 50505040, 50504140, 50414140, 41414140, and so on for 14 more. Some of these modes are balanced as to equal on and OFF intervals; the others are not as much unbalanced, and therefore inadmissible, as the mode 5040 for frequency ratio 9. The transient output will have somewhat of a "jitter," but as this is a non linear system, purely sinusoidal oscillations are not to be expected.

Fortunately, the admissibility of these additional modes of operation, of which there are an infinite number, does not alter the design criterion given previously. However, the phase graph for ϕ_N becomes an infinite number of lines, one at each frequency. If we let $H=C$ when CT is odd, and $C/2$ when CT is even, then in general the boundaries of the various families of ϕ_N lines are

$$-(\pi-\phi_{\max})=-\left(\pi-\frac{\theta}{2}+\frac{\theta}{4H}\right) \tag{53}$$

$$-(\pi-\phi_{\min})=-\left(\pi-\frac{\theta}{2}-\frac{\theta}{4H}\right) \tag{54}$$

It is noteworthy that all phase ranges still are centered about the line $-(\pi-\theta/2)$, and that they grow smaller, approaching a point on the center line, as H (and C) approach infinity. The magnitude graph for $1/|N|$ also has ranges at every frequency at which a ϕ_N line is crossed by ϕ_G. These ranges grow in number and decrease in extent as H and C increase.

It can now be seen that, as a stable system with a principal time constant considerably greater than T_s grows unstable as the gain is increased or the inactive zone decreased, the first conditions for which the Barkhausen criterion is satisfied will be for even-integer frequency ratios T. However, for high gains, or for systems in which ϕ_G rapidly penetrates the potentially unstable phase region, some of these additional modes will satisfy the criterion for continuous oscillation.

Each mode of operation satisfying the criterion represents a sort of limit cycle for sustained oscillations. Some of these are convergent (stable), some divergent (unstable), and some possibly divergent-convergent (half-stable).[6] Generally a limit cycle is defined in a manner such that as the signal amplitude increases, convergent and divergent limit cycles appear alternately. Also in the case of true limit cycles, all initial conditions having amplitudes between given adjacent convergent and divergent limit cycles will result in a steady-state oscillation in the particular convergent mode. The term "sort of limit cycle" was used here for the sampling servo because neither of these restrictions apply to the limit cycles encountered. This phenomenon has been remarked upon by A. E. Bailey,[7] and was substantiated by the writer.[1] Apparently Dr. Chow has considered only the possibility of convergent limit cycles, which will be the first to appear as the gain is increased beyond the stable limit. Divergent and divergent-convergent limit cycles are difficult to locate by computer or transient analysis because it seems thus far to be impossible to predict the right initial conditions to produce them. Even the establishment of a complete set of convergent limit cycles on a high-gain system with many unstable points was difficult on the analogue computer I used, partly again because of difficulty in finding the right initial conditions and partly because small occasional imperfections in the sampler-hold-contactor operation caused such frequent "mode hopping" that it was often impossible to tell whether a given mode was truly convergent or whether it was merely transient. It is felt that the describing function analysis may prove a useful tool for locating and studying the occurrence of these phenomena.

As far as design is concerned, the presence of any of the three types of limit cycles indicates an unstable system as defined by Dr. Chow. However, when an oscillatory-control servo[8] is desired, the frequency-response method may be used to provide a design giving a single intersection. Such an intersection will necessarily represent a convergent limit cycle of small amplitude, as would be desired.

Accuracy of Frequency-Response Method

From transient studies it was determined that a slow, nonsinusoidal convergent limit cycle of amplitude barely exceeding the inactive zone may occur, which is not predicted by the Barkhausen criterion, for systems near instability. The reason for the nonsinusoidal form of the oscillation is that the servo is "coasting" within the dead zone most of the time. The reason that this peculiar mode does not show up in the frequency-response analysis is that the fundamental component of the error signal does not exceed the inactive zone. Adequate gain margin in design will prevent the occurrence of such an instability.

Another point of interest bearing on accuracy can be found from harmonic analysis of the contactor output for all possible modes. In brief, it is found that the fundamental component is largest relative to the harmonics for even-integer T, and for the middle modes for a given even-integer T. Thus the accuracy is best just where it is most needed for the design of stable or oscillating-control servomechanisms.

Lastly, I would like to point out that nonidealized contactor characteristics such as asymmetrical inactive zone, hysteresis, and saturation with a linear zone, and clamper characteristics which do not hold the signal for the complete period T_1, can be handled by these methods. Hysteresis, for example, produces new modes having ϕ_N lines projecting above the boundary $-(\pi-\theta)$ and therefore makes the system less stable. However, in terms of an analysis already so complicated, it is difficult to generalize when an additional variable involving the inactive zone is introduced. For a particular system the work involved would not be prohibitive. On the other hand, a complication of $KG(p)$ adds substantially no difficulty to the method, especially if logarithmic magnitude and frequency scales are used.

References

1. Design Criterion for Stability of Sampled-Data On-Off Servomechanisms, **Frederick A. Russell.** *Thesis*, Columbia University, New York, N. Y., June 1953.

2. Consideration of Off-On-Modulated Reversing Clutch Servo Systems, **T. R. Stuelpnagel, J. P. Dallas.** *AIEE Transactions*, vol. 7, pt. II, 1952, pp. 466–10.

3. The Dynamics of Automatic Controls (book), **R. C. Oldenbourg, H. Sartorius.** The American Society of Mechanical Engineers, New York, N. Y., 1948, chapter V-A, p. 180.

4. The Analysis of Sampled-Data Systems, **J. R. Ragazzini, L. A. Zadeh.** *AIEE Transactions*, vol. 71, pt. II, 1952, pp. 225–34.

5. Theory of Servomechanisms (book), **H. M. James, N. B. Nichols, R. S. Phillips.** McGraw-Hill Book Company, Inc., New York, N. Y., 1947, p. 236 et seq.

6. See reference 11 of the paper, p. 62 et seq.

7. Characteristics of Sampling Servo Systems, **C. Holt Smith, D. F. Lawden, A. E. Bailey.** *Proceedings*, DSIR Conference on Automatic Control, Cranfield, England, July 1951. See reply to discussion by A. E. Bailey. (Also published in book form as Automatic and Manual Control edited by **A. Tustin.** Butterworths Scientific Publications, London, England, 1952, p. 407.)

8. Fundamental Theory of Servomechanisms (book), **L. A. MacColl.** D. Van Nostrand Company, Inc., New York, N. Y., 1945, p. 73 et seq.

C. K. Chow: I wish to thank Dr. Russell for his discussion. It is very interesting to see the construction of the critical regions on the Bode diagram, in which the regions are represented by a family of vertical lines on the phase and magnitude plots. The $1/|N|$ ranges are a family of infinite sets of lines, one set of lines being constructed for each depth of penetration of ϕ_G into the ϕ_N lines.

I wish to emphasize that the following conclusions are reached through rigorous proofs, and that they are *not* assumptions made to simplify the analysis:

1. The dimensionless period of self-sustained oscillation of the servo system is an integer, and the corresponding sampler, clamper, and contactor outputs are all periodic and have the same period as the system output.
2. One or more integration factors in $g(s)$ will assure that the positive and negative contactor ON intervals must be of equal duration.

They are true no matter whether the error signal is symmetrical or not (that is, the average component of the error signal is zero or not).

For instance, in the example of Dr. Russell, mode 3031, which is essentially the same mode as listed in Table V, mode 2, consists of equal positive and negative ON intervals: namely, three. The value of δ for this particular mode is, therefore, three, and the mode is designated as $T=7$ and $\delta=3$ in the paper. In the terminology of Dr. Russell, the mode as illustrated in Fig. 18 will be designated as 1111, while in the paper it is designated by $T=4$ and $\delta=1$. The mode number designation δ is adequate for both odd ($T=4Q+1$ or $T=4Q+3$) and even periods, whenever $\lambda\geq 1$.

The critical regions as derived in the paper are applicable to the case of asymmetrical error signals as well as the case of symmetrical ones. (Only $\lambda\geq 1$ is under consideration.) In fact, the modes as given in Table II, no. 3, Table IV, case 2, no. 2, and Table V, no. 2, all represent the cases having asymmetrical error signals. The reason for this is that the critical region of any values of δ and T for the asymmetrical case is smaller than and contained in the critical region of the same values of δ and T for the symmetrical case. It can be proved as follows:

Due to the limited space, only the proof for the case of $T=4Q$ and $\delta=2i$, where $1\leq i\leq Q-1$, is given here. Let the asymmetrical error signal fed to the nonlinear element be

$$A+\epsilon=A+|\epsilon|\sin(u\tau+\phi) \tag{55}$$

where A is the average component of the error signal and ϵ is the sinusoidal component of the error signal. Without loss of generality, the constant A may be assumed as positive.

In parallel to the derivation of equation 27, it is obtained that to ensure $\delta=2i$ the conditions on ϵ are

$$\begin{aligned}&|\epsilon|\sin[(Q-i)\theta+|\phi|]+A>\Delta>|\epsilon|\times\\&\qquad\sin[(Q+i)\theta+|\phi|]+A\\&|\epsilon|\sin[(Q-i)\theta+|\phi|]-A>\Delta>|\epsilon|\times\\&\qquad\sin[(Q+i)\theta+|\phi|]-A\end{aligned} \tag{56}$$

Therefore the limiting values of $|\epsilon|$ are

$$|\epsilon|_{\min}=\frac{\Delta+A}{\sin[(Q-i)\theta+|\phi|]}=\frac{\Delta+A}{\cos\left(\frac{\delta}{2}\theta-|\phi|\right)} \tag{57A}$$

$$|\epsilon|_{\max}=\frac{\Delta-A}{\sin[(Q+i)\theta+|\phi|]}=\frac{\Delta-A}{\cos\left(\frac{\delta}{2}\theta+|\phi|\right)} \tag{57B}$$

Comparison of equations 57(A) and (B) with equation 28 reveals that the values of $|\epsilon|_{\min}$ and $|\epsilon|_{\max}$ of the asymmetrical case are respectively greater and smaller than the values of $|\epsilon|_{\min}$ and $|\epsilon|_{\max}$ of the symmetrical case. Consequently the critical region of the asymmetrical case is contained in that of the symmetrical case.

It is interesting to see from equation 57(B) that whenever $A\geq\Delta$, $|\epsilon|_{\max}\leq 0$. This is, of course, not admissible; in consequence, it may be concluded that no mode of self-sustained oscillation having an average component greater than Δ and characterized by $T=4Q$ and $\delta=2i$ $(1\leq i\leq Q-1)$ can occur.

Considerations in Applying Rectifiers As a Power Supply for Hot Strip Mills

G. M. ZINS
ASSOCIATE MEMBER AIEE

E. J. CHAM
MEMBER AIEE

IN RECENT years the trend has been to apply ignitron rectifiers as a power supply for continuous hot strip mill main drive service. As in numerous other fields, the versatile ignitron is again demonstrating its many inherent economic advantages over conventional rotating machines and proving its rugged dependability for main drive service.

By capitalizing on the inherent desirable features of ignitron rectifiers and by improving the design of the ignitron tubes and their associated control and auxiliaries, the rectifier power supply has been molded to provide very nearly the ultimate in modern hot strip mill drives. This is evidenced by the installation at the Fairless Works of the United States Steel Corporation at Morrisville, Pa. New techniques recently developed will permit further simplification of the rectifier control and auxiliaries.

This paper has a twofold purpose. First, it will describe the ignitron power supply for main drive service in the continuous hot strip mill recently installed at the Fairless Works; and second, it will present a few of the more important considerations involved in applying ignitron rectifiers to variable voltage drives of this nature.

Fairless Works

Fig. 1 shows the ignitron installation in the motor room of the 80-inch continuous hot strip mill of the Fairless Works. The ignitron assembly is shown on the left and the anode breaker cubicles and auxiliary cubicles to the right of the operator. The entire hot strip mill consists of a scale breaker, three roughing stands, and six finishing stands. Fig. 2 shows a simplified 3-line diagram of the power equipment of the six finishing stands.

It should be noted from Fig. 2 that the three d-c busses are independent and that the power equipment supplying each bus is identical. Each d-c bus is energized from two 12-tube ignitron rectifier assemblies. Each ignitron assembly has a continuous rating of 4,285 kw at a maximum voltage of 750 volts d-c. Each ignitron assembly is energized from a 12-phase rectifier transformer through two 6-pole high-speed anode circuit breakers. The rectifier transformers have their primaries connected delta and their secondaries connected quadruple zigzag. Taps are provided on the primary of each transformer to give maximum d-c bus voltages of 700, 650, 530, 430, and 300 volts at 200-per-cent load. The transformers are oil-insulated and self-cooled for outdoor service. A single a-c circuit breaker connects 13.8-kv 60-cycle power to the two rectifier transformers associated with each d-c bus.

The d-c output of each rectifier assembly is connected to the d-c bus through a single-pole, semi-high-speed, air circuit breaker. The no. 1 d-c bus supplies power for two 5,000-horsepower (hp) d-c mill motors which drive stands 5 and 6 of the hot strip mill. Bus no. 2 also supplies two 5,000-hp mill motors for stands 7 and 8. Bus no. 3 supplies a 5,000-hp motor for stand 9 and a 4,000-hp motor for stand 10. Each mill motor is connected to its respective bus through two single-pole, semi-high-speed, air circuit breakers. A dynamic braking resistor is connected to each mill motor through a single-pole, semi-high-speed, air circuit breaker.

A MAGAMP voltage regulator is supplied for each d-c bus. A calibrated d-c reference voltage, which is common to all three busses, supplies the operating reference voltage for all three MAGAMPS.

Normal start up of the ignitron rectifiers is from the duplex switchboard in the motor room with a single control switch for each d-c bus section. The time required for the start up of a bus section is only a few seconds. After start up, the rectifier equipment is left unattended and all control is from the mill pulpit. In the mill pulpit a single control switch brings the voltage of the three busses from 30 volts to maximum in about 15 seconds. As previously mentioned, the maximum direct voltage is determined by the voltage tap of the rectifier transformer.

The protective devices of the rectifier installation are arranged for a case of trouble to either lock the rectifier out of service or stop the rectifier with automatic restarting allowed when the trouble no longer exists.

The following listing indicates the trouble conditions which initiate lockout and those which merely stop the rectifier with automatic starting allowed.

SHUTDOWN AND LOCKOUT

1. Repeated arc backs within a definite time interval.
2. A-c overcurrent.
3. Low oil level on the rectifier transformer.
4. Loss of excitation to the ignitrons.
5. Rectifier transformer high tank pressure.
6. Ignitron cooling water low pressure.
7. Ignitron rectifier high temperature
8. Mercury vapor vacuum pump high temperature.

SHUTDOWN WITH AUTOMATIC RESTARTING

1. Single phase, reverse phase, or low auxiliary alternating voltage.
2. Loss of ignitron tank vacuum.
3. Rectifier transformer high temperature.

All conditions which result in a shutdown are annunciated.

Protective circuits for ignitron rectifier

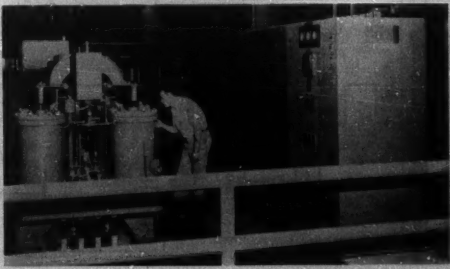


Fig. 1. Fairless Works hot strip mill motor room

Paper 54-107, recommended by the AIEE Mining and Metal Industry Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing November 30, 1953.

G. M. ZINS and E. J. CHAM are with Westinghouse Electric Corporation, East Pittsburgh, Pa.

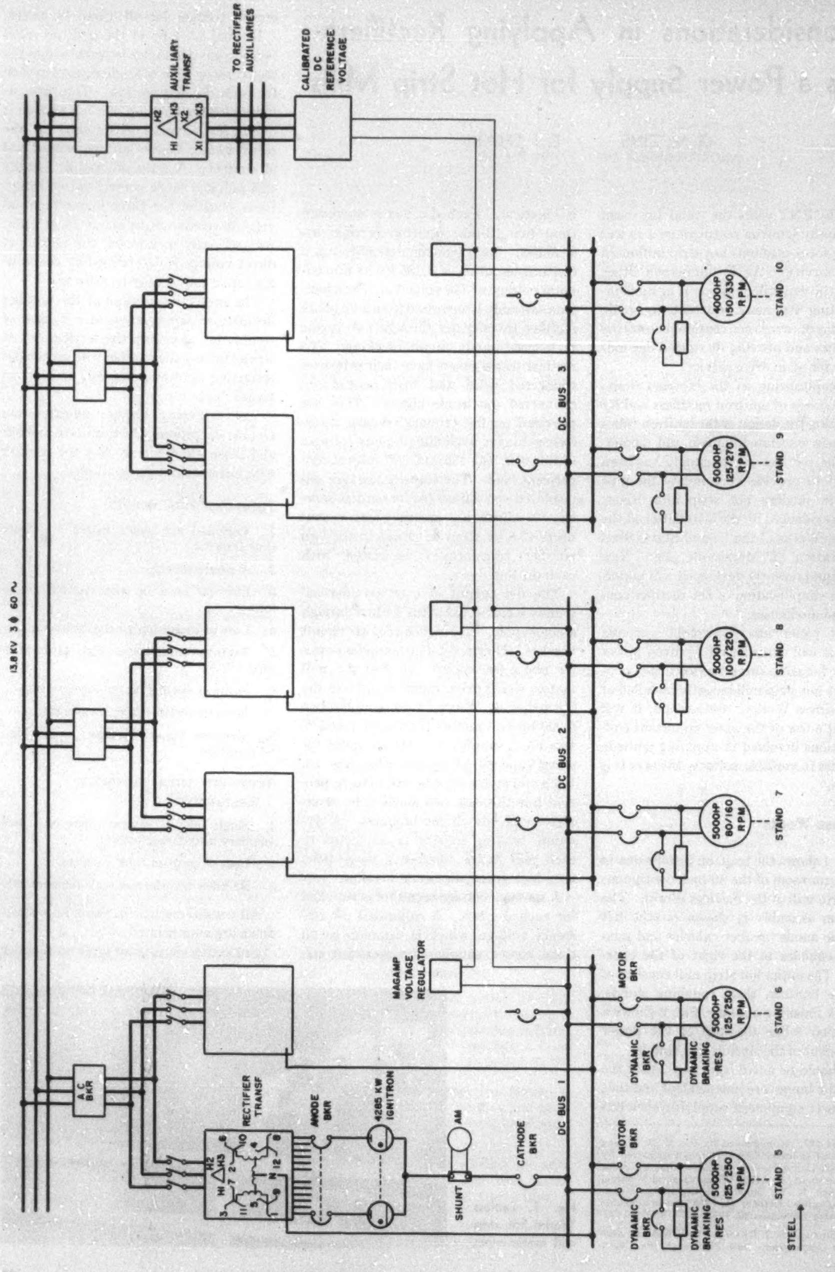

13.8 KV 3Φ 60~
AUXILIARY TRANSF
H1
H2
H3
X1
X2
X3
TO RECTIFIER AUXILIARIES
CALIBRATED DC REFERENCE VOLTAGE
A C BKR
RECTIFIER TRANSF
ANODE BKR
4285 KW IGNITRON
AM
SHUNT
CATHODE BKR
MAGAMP VOLTAGE REGULATOR
DC BUS 1
DC BUS 2
DC BUS 3
MOTOR BKR
DYNAMIC BKR
DYNAMIC BRAKING RES
5000HP 125/250 RPM
STAND 5
5000HP 125/250 RPM
STAND 6
5000HP 80/160 RPM
STAND 7
5000HP 100/220 RPM
STAND 8
5000HP 125/270 RPM
STAND 9
4000HP 150/330 RPM
STAND 10
STEEL

Fig. 2. Elementary diagram of power supply for Fairless Works hot strip mill

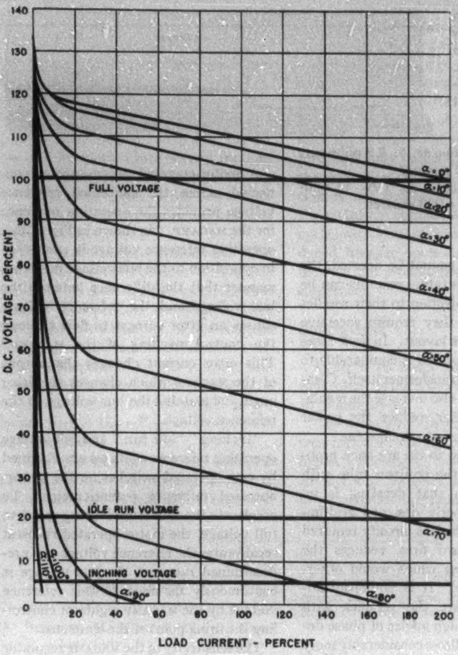


Fig. 3. Typical rectifier regulation characteristic curve

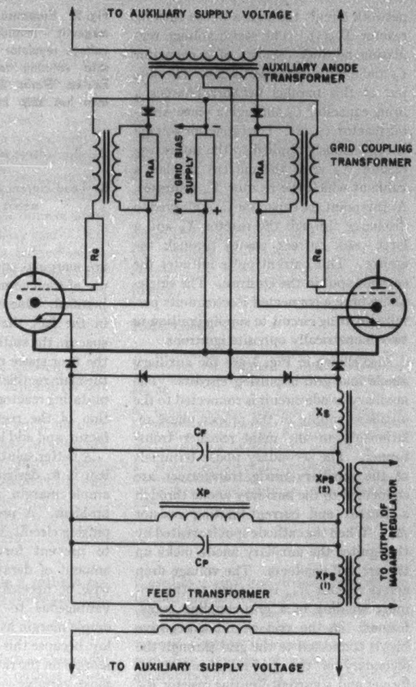

TO AUXILIARY SUPPLY VOLTAGE
AUXILIARY ANODE TRANSFORMER
GRID COUPLING TRANSFORMER
TO GRID BIAS SUPPLY
-
+
RAA
RAA
RG
RG
XS
CF
XPS (2)
XP
CP
XPS (1)
TO OUTPUT OF MAGAMP REGULATOR
FEED TRANSFORMER
TO AUXILIARY SUPPLY VOLTAGE

Fig. 4. Ignitron rectifier grid-impulsing circuit

high temperature, mercury vapor vacuum pump high temperature, and ignitron cooling water low pressure are connected to a lockout relay and also to a shutdown relay through a transfer timing relay. This arrangement is provided so that the operation of these protective devices results in shutdown and lockout if the rectifier has been in operation at least 10 minutes, but only prevents starting for 10 minutes if the rectifier has been shut down to loss of or low auxiliary alternating voltage. This is done to prevent lockout during a shutdown that resulted from the loss of a-c power because the latent heat of the rectifier units may operate the temperature devices. This same arrangement is also provided for the cooling water low pressure switch since it may operate because of low voltage on the pump motor.

Major Considerations

There are many considerations involved in designing any ignitron rectifier installation. It is within the scope of this paper to consider only matters deemed particularly unique or important to variable voltage drives. These considerations are voltage control, voltage regulation, and mill control.

VOLTAGE CONTROL

Voltage control on an ignitron rectifier is accomplished by phase control of the ignitor pulse. For applications which require reduced voltages, of 5 to 10 per cent of rated voltage, it may be necessary to phase-delay the ignitors as much as 120 degrees. The amount of phase delay is limited by the rectifier transformer interphase design.

Fig. 3 shows a typical rectifier regulation characteristic. Regulation curves are shown for angles of delay (α) from zero to 110 degrees. The light-load voltage rise has been exaggerated for clarity. It should be noted from the curves that approximately 110 degrees delay is necessary to obtain the desired value of "inching" voltage at light load.

Phase delay is obtained with a nonlinear type of reactor firing circuit which has a 120-degree phase-shift range. The phase angle of firing is varied by the control current flowing in the d-c windings of the phase-shifting reactors. This control current is supplied by the MAGAMP voltage regulator.

Operation at high angles of delay presents two problems which require consideration. With an increase in delay the probability of both forward fires and arc backs is increased. Special means are required to prevent this.

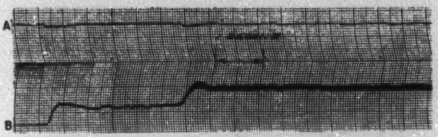
The problem of preventing premature conduction or forward fires is readily solved by applying a negative blocking voltage to the grid of the ignitron, as shown in Fig. 4. Fig. 4 shows a simplified schematic diagram of the excitation circuits. These circuits include the firing circuit, auxiliary anode circuit, and grid impulsing circuit. The feed transformer supplies the operating voltage of the firing circuit in the correct phase relationship to the main rectifier transformer. The secondary of the feed transformer is connected to a static voltage regulating network circuit through a phase shifting reactor $X_{PS}(1)$. The static voltage regulating network consists of a capacitor C_P and a saturating reactor X_P. The output of this parallel network charges a firing capacitor C_F through a phase-shifting reactor $X_{PS}(2)$. The capacitor voltage is continuously applied to the saturating reactor X_S until the current reaches a value at which the reactor X_S saturates. At this point the capacitor C_F is allowed to discharge through the reactor X_S and a large peak current passes through the ignitor. This current pulse initiates the cathode spot in the ignitron. The single-phase bridge-connected Rectox units permit one firing circuit to supply ignition to two diametrically opposite ignitrons.

Also shown in Fig. 4 are the auxiliary anode and grid-impulsing circuits. The auxiliary-anode circuit is connected to the auxiliary supply in the proper phase relationship to the main rectifier transformer. The secondary phase terminals of the auxiliary-anode transformer are connected to the auxiliary anode through a Rectox and current-limiting resistor R_{AA}. When the cathode spot is created by the ignitor the auxiliary anode picks up the arc and conducts. The voltage drop across resistor R_{AA} is applied to the primary winding of a grid-coupling transformer. In the grid circuit a negative bias is connected to the grid through the secondary of the grid-impulsing transformer and a current-limiting resistor R_G. The secondary winding of the grid-impulsing transformer is connected to oppose and overcome the negative bias on the grid. The grid is therefore held negative until it is pulsed positive when the ignitor is fired and conduction is desired.

The second problem is that of increased arc-back probability at the higher angles of delay. It has been shown that the frequency of arc back is approximately proportional to the tenth power of arc-back factor.[1] The arc-back factor is proportional to the product of the initial inverse voltage on the tube and the final rate of decay of current in the tube at the end of the conduction cycle. The initial inverse voltage increases with higher angles of delay. The commutating angle decreases with higher angles of delay.[2] As the commutating angle decreases, the rate of decay of current increases in the anode being relieved at the end of its conduction period. The problem of increased arc backs at higher angles of delay can be solved by one of two generally accepted methods.

One method that can be used to reduce the arc-back factor is to employ commutating reactors.[1] Where tube ratings are marginal and cannot be derated, the use of such commutating reactors can be justified. One objection to their use lies in the fact that they require excessive space in the station layout. In some cases the floor space required is comparable to the main rectifier transformer itself. Commutating reactors also increase the regulation of the rectifier, reduce the power factor, and add losses to the system.

Fig. 5. Response of magnetic - amplifier voltage regulator to strip entering roll. Fairless Works 80-inch hot strip mill



A—Bus voltage section no. 1; 2.8 millimeters = 1 per cent
B—Load current rectifier no. 1-A; 14 millimeters = 4,000 amperes

A better solution to the arc-back problem is to design the ignitron tube with ample margin so that derating is no problem. A properly designed grid-impulsing circuit, which is already required to prevent forward fires, reduces the amount of derating which would otherwise be necessary. It is obviously advantageous to rate the rectifier with ample margin at high angles of phase delay, because this allows considerably more margin on the rating with no delay. Since most variable-voltage applications operate the great majority of the time with very little delay, this gives a greater over-all margin to the equipment. Other advantages to this method are that no additional space is required, power factor and efficiency are not decreased, and the regulation is not increased.

VOLTAGE REGULATION

Referring to Fig. 3, it can be seen that the slope of the regulation curves becomes quite steep at light loads. When a strip enters a stand the load on the rectifier may suddenly be increased from idling load to as much as 200-per-cent load. Because of the characteristic shape of the regulation curves and the type of load applied, the voltage regulator must have a high speed of response and a high degree of stability. In order to permit the operator to maintain close speed adjustment of the mill motors by means of the vernier field rheostats, it is essential that the bus voltage be held with very narrow tolerances. This problem becomes even more acute if a split-bus arrangement is employed. Static voltage regulators of the magnetic-amplifier type are particularly well suited to meet these requirements.

A motor-operated potentiometer connected across the calibrated reference voltage provides the operating reference for the MAGAMP. As shown in Fig. 2, this operating reference voltage is connected in opposition to the bus voltage in such a manner that the difference between the bus voltage and the reference voltage causes an error current to flow through the control winding of the MAGAMP.[3] This error current changes the output of the MAGAMP which changes the firing angle and matches the bus voltage to the reference voltage.

"Inching," "idle run," and full voltage operating reference voltages are obtained by cam-operated switches on the motor-operated reference potentiometer. To accelerate the mill from inching speed to full voltage, the motor-operated rheostat recalibrates the reference voltage at a predetermined rate and the bus voltage is continuously matched to this reference voltage by the MAGAMP regulator controlling the firing point of the ignitrons.

The sensitivity of the MAGAMP regulator is ±0.25 per cent and the speed of response is better than 0.25 second for full correction. Fig. 5 shows a typical response characteristic of the MAGAMP to a strip entering the roll. It can be seen that the voltage is completely recovered within this time.

Rated full load voltage is normally required from 170 to 200 per cent of rated load for steel mill applications; however, any desired regulation characteristic is readily obtained by proper co-ordination between the design of the rectifier transformer, excitation circuits, and voltage regulator.

MILL POWER CONTROL

In a hot strip finishing train, roll speeds require a high degree of synchronization between stands to prevent tearing of the strip being rolled. As the strip enters successive stands, high impact loads are encountered and these present a severe test of voltage regulation of the d-c power system. For successful operation it is necessary that the voltage applied to all mill motors be the same regardless of their loading. Thus, it has been usual practice to use a single d-c bus as a power supply for the entire finishing train.

It is well known that a MAGAMP voltage

regulator has a high degree of sensitivity. A MAGAMP voltage regulator in combination with an ignitron rectifier results in a d-c power system with a high speed of response and high degree of sensitivity. When these two features are utilized in a power system for a hot strip mill, the d-c bus can be divided into several sections instead of the customary single bus required of other types of d-c power supplies. A split-bus system of this type has several advantages.

A d-c power supply with several independent bus sections permits considerable flexibility to meet special operating conditions. Independent bus sections have the important advantage of allowing the operation of mill motors at different voltages. This permits a wider range of mill motor speeds than would normally be available. Wider ranges of mill motor speeds permit the rolling of special steel and steel sizes.

Initial cost of a split d-c bus system is usually less since interrupting capacity of d-c breakers can be reduced. This allows the use of small-capacity breakers without reducing mill capacity.

A d-c power supply utilizing ignitron rectifiers can considerably reduce breaker operations in the normal starting and stopping operations of the mill. To reduce the output voltage of an ignitron to zero, it is only necessary to short-circuit the excitation circuit. A breaker operation is not necessary. Short-circuiting the excitation circuit can be done in a few cycles with a simple auxiliary relay. Reduced breaker operations reduce maintenance and prolong the useful life of the equipment.

Conclusions

Ignitron rectifiers have proved to be well suited as main power supplies for hot strip mill drives.

The usual advantages of ignitrons over conventional rotating power supplies are higher efficiency, lower initial costs, simpler and lower installation costs, and lower maintenance expense.

Ignitron rectifiers are well suited for steel mill control operations since the starting and stopping sequences are quite simple. Voltage control with ignitrons is also simplified since the regulator need only carry a small current to control bias voltage. Breaker operations can be reduced by the use of ignitron short-circuiting, which is another advantage over rotating machines.

These factors, coupled with the proved reliability of ignitrons for steel mill service, indicate that ignitrons will be the conventional hot strip mill drive of the future.

References

1. NONLINEAR COMMUTATING REACTORS FOR RECTIFIERS, A. Schmidt. *AIEE Transactions (Electrical Engineering)*, vol. 65, Oct. 1946, pp. 654–56.

2. HIGH VOLTAGE IGNITRON RECTIFIERS AND INVERTERS FOR RAILROAD SERVICE, J. L. Boyer, C. G. Hagensick. *AIEE Transactions (Electrical Engineering)*, vol. 65, July 1946, pp. 463–70.

3. MAGNETIC-AMPLIFIER APPLICATIONS IN D-C CONVERSION STATIONS, W. A. Derr, E. J. Cham. *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 220–29.

◆

Discussion

A. Schmidt, Jr. (General Electric Company, Schenectady, N. Y.): The authors have described a system which represents an important forward step in reduction of available short-circuit current on a d-c bus. Still further reduction is achieved by the connection of a single rectifier to each of their individual motors. The reliability of present-day rectifiers as evidenced by their performance record justifies such a step.

By contrast, the reliability of rectifiers was not established when the first proposal for application to hot-strip mill service was entertained nearly 15 years ago. Anode breakers had not been developed at this time, and the need for uninterrupted service with voltage reduction of 10 to 30 per cent by phase control prompted the use of commutating reactors to secure the utmost in reliability. The high grade of performance of the earlier installations, even under extreme overload conditions, has been an important factor in the general acceptance of rectifiers for main roll drive by the steel industry. The subsequent development of rectifiers with improved performance and the use of anode breakers have limited the need for commutating reactors, and their use is being discontinued for applications of this nature.

The authors have stated that the subject rectifiers are capable of maintaining rated voltage up to 200 per cent load. We are interested to know whether this capability is present only at rated a-c voltage, or at reduced voltage. The rectifiers described in the discussion by Mr. Larson, and other hot strip mill rectifiers of the same manufacture, are designed to deliver rated voltage at loads up to 175 per cent of the rectifier rating with the supply voltage at 95 per cent of its rated value. This is a realistic appraisal of voltage and load conditions that frequently arise in service.

H. E. Larson (General Electric Company, Schenectady, N. Y.): The authors are to be complimented for describing so clearly equipment that solves the critical problems of voltage control and voltage regulation encountered when using rectifiers to supply a hot-strip mill installation such as at Fairless Works, using a permanently split bus and magnetic-amplifier type of voltage regulators. In commenting on this paper it will be my aim to stress further the importance of the problems discussed by the authors by touching upon the features of equipment now being built for other hot-strip mills and also to mention some other considerations not mentioned in the paper.

The experience gained on a 56-inch hot-strip mill in the Pittsburgh, Pa., area which has been in successful operation for about $1^{1}/_{2}$ years, using one generator for each of six stand motors without bus connection between stands and AMPLIDYNE-type voltage regulators, has demonstrated that hot-strip mills can be so operated without introducing undesirable tension and looping effects on the strip under high-speed threading and rolling conditions. Rectifiers having regulating systems using magnetic amplifiers such as the AMPLISTAT or MAGAMP in general have higher rates of response and therefore should have some slight advantage over systems using generators.

One 66-inch hot-strip mill now being modernized and speeded up will employ one 1,500-kw generator and one rectifier for each finishing stand motor without bus connection between stands. Rectifiers vary in size from 2,500 to 3,000 kw. An AMPLISTAT voltage regulator for each stand system is used to control load division between the rectifier and generator, as well as to control voltage applied to the stand motor using phase control in accordance with the dictates of a voltage reference common to the regulating systems of each stand. The generators will normally be used for starting the stand motors under no load but the rectifiers are also equipped with 100-per-cent phase control for mill-starting purposes.

Another 60-inch hot-strip mill will employ one rectifier only, varying in size from 2,800 kw to 4,000 kw for each finishing stand motor without bus connection between stands. Here the AMPLISTAT voltage regulator for each stand system controls voltage applied to the stand motor using a combination of phase control and transformer tap changing under load for the continuous operating voltage range. For short-time starting periods 100-per-cent phase control is employed.

In each case the design of the AMPLISTAT regulating system is based on results of studies made on analogue computors which show system adjustments necessary for stable regulating systems having acceptable errors of the magnitude shown in the high-speed recording chart of Fig. 5.

Dynamic braking applied with contactors is a feature of these new mill installations. Dynamic braking is used not only for removing energy from the motors and stands during emergency stopping. When a rectifier only supplies a stand motor it is also used as a means for limiting overvoltage when rapidly strengthening the field of a stand motor, which sometimes occurs when changing the rolling speed schedule. Normally, the possibility of overvoltage by field strengthening occurs only at no load without strip in the mill.

The use of an individual power system for each stand reduces to a minimum the short-circuit currents that can flow through motor series windings and breakers. It also reduces the number of circuit breakers as only one semihigh speed circuit breaker is required for the rectifier-motor loop circuit. As pointed out in the paper, maximum rolling mill flexibility also results when speed changes are obtained by varying the voltage on each stand permitting the use of maximum field strength in the motors for maximum torque output.

It is clear from the author's description that the Fairless Works installation uses pumped type ignitron rectifiers. It should be of interest that the experience with pumpless, permanently sealed (not sealed tube) rectifier tanks has now proceeded to the point where 48,000 kw of capacity for steel mill main roll drives is now installed or on order.

G. M. Zins and **E. J. Cham**: The discussions are informative additions to the paper, and are appreciated. Mr. Schmidt suggests using a single rectifier connected to each individual motor. Although present-day performance of ignitron rectifiers would justify consideration of using one rectifier per mill motor, other problems would not make such a system very desirable. For example, if a 6-tube rectifier were used to supply a single-mill motor, an arc back on one tube would reduce the loading capacity of the rectifier installation to such a degree so as to prohibit the continuing of the rolling operation until the fault had cleared and the anode breaker reclosed. Since usual practice requires a time interval between reclosures, a severe limit would be imposed on the system. Obviously as the number of tubes used in an installation is increased, a single-tube outage is less important. The application of a single rectifier to a single-mill motor should be considered on an individual application basis, giving the various factors due consideration rather than considering this as a general practice.

In reply to Mr. Schmidt's question, present standard National Electrical Manufacturers Association standards require that 100 per cent direct voltage should be available at specified currents with 95 per cent of rated alternating voltage available. However, it should be noted that unnecessary compensation in the rectifier transformer penalizes the user since under normal alternating voltage conditions, more phase delay is required than would normally be required, resulting in a lower power factor on the system. In general, flat regulation can be provided for rated voltage or any desired overload, with or without compensated transformers.

In regard to Mr. Larson's comments, we still recommend the pumped tube for heavy-duty steel mill service, since long range operation results in economies over the sealed or pumpless type of tube. Where specific conditions justify the use of a sealed tube design, these are available in high-capacity ratings for steel mill service.

Fundamentals of Flashing of Diesel-Electric Motors and Generators

C. A. ATWELL
FELLOW AIEE

FLASHING is one of the important considerations in the manufacture and operation of electrical equipment for diesel-electric locomotives. Its importance has increased with the placing in service of a large number of high-speed and high-powered road locomotives during the past six or seven years. Examination of a large number of motors and generators of different manufacture on high-speed road locomotives has indicated that the machine which shows no evidence of flashing is an exception. Flashing may occur without any appreciable damage to a motor or generator or any interruption of service. Flashing, on the other hand, may be so serious that the locomotive cannot continue in service or that a motor or generator has to be removed for repair.

It is the purpose of this paper to examine some of the fundamental causes of flashing from both design and operating standpoints, so that better understanding may be had of how to minimize flashing and its effects.

What Is Flashing?

Flashing has been considered by many persons as mysterious, because like lightning, it occurs suddenly and often when least expected. There are several kinds of flashing. It is for this reason that the term has been used in the subject of this paper rather than "flashover." An arc or flash may occur where commutator bars are leaving the trailing edge of the brush. This, occurring suddenly, is certainly a visible flash even if it does not extend to any other part of the machine. Such an arc or flash at the brush edge may appear to be quite severe without producing an actual flashover between brush holders. A flash of light around the commutator commonly known as "ring-fire" may also occur. Ring-fire is the result of small arcs between one or more pairs of adjacent commutator bars. These arcs between bars are usually formed as the spaces between adjacent bars short-circuited by the brush leave the brush, and may continue for several revolutions. In cases of exceptionally high maximum volts per bar resulting from high voltage and a distorted field form, these arcs may form while the bars are between brushes. The apparent continuous ring of fire around the commutator is usually an optical illusion caused by the high peripheral commutator speed while these small arcs exist between certain bars. There also may be a flash or arc from the commutator or brushes to some adjacent grounded part of the machine as a result of creepage over an insulating surface.

The two most common kinds of flashing are the arcing at the edge of the brush and ring-fire. Neither one of these is usually very destructive in itself, but either one may develop into an arc extending over the surface of the commutator from brush holder to brush holder. This type of flashing is defined as a flashover and can be quite destructive if allowed to continue for even a fraction of a second.

Arcing at the brush edge is caused by a current surge, and actually is instantaneous bad commutation. Any arcing of this kind creates vaporized conducting material, and if the heat of the arc is sufficient to vaporize enough material and ionize the surrounding air, the arc will grow in size and may result in a complete flashover. The instantaneous condition of bad commutation does not mean that the motor or generator is a poor commutating machine. Present-day diesel-electric motors and generators usually have excellent commutating ability over their entire range of operation under steady-state conditions. The instantaneous bad commutation which initiates a flashover is caused by a sudden surge of armature current. Since there is a time lag in the commutating field magnetic circuit, the voltage induced in the coils short-circuited by the brush is not fully

---

Paper 54-45, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted Oct. 20, 1953; made available for printing Nov. 17, 1953.

C. A. ATWELL is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

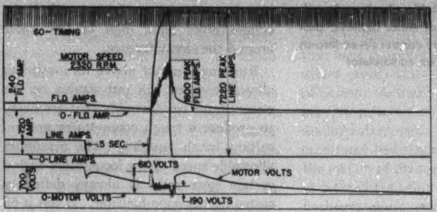


Fig. 1 (left). Tracing of oscillogram of motor flashover produced by current interruption method

compensated for by the commutating field for a short period of time, resulting in instantaneous bad commutation.

The ring-fire kind of flash also may result in a complete flashover by an increase in the number of small arcs between bars until a heavy arc extends from brush holder to brush holder.

Oscillographic analysis of flashovers has proven valuable in determining the current and voltage variations that occur. Such analysis on locomotives in service, however, has not proven very fruitful because of the difficulty of obtaining flashovers when wanted. They seem to occur always at times other than when the oscillograph equipment is set up on the locomotive. Oscillographic analysis has been more profitable on the test floor where the severity of conditions can be increased until flashover does occur.

Fig. 1 is a tracing of an oscillograph of a flashover of a series railway motor produced by the ordinary circuit interruption and reclosing method. Connections were as in Fig. 2. The interruption was for 1/2 second. The conditions prior to the interruption of the circuit correspond to running at maximum speed in weak field. The voltage was 20 per cent higher than is ever used in service at this field strength. To produce the flashover, this higher voltage was required even with the severe condition of circuit interruption for 1/2 second. When the circuit was reclosed, the line current rose very sharply and the flashover started in approximately 1/60 second. This is indicated by the beginning of the rapid oscillation of the field current and voltage traces caused by changes in the arc. The line current and field current continued to rise to peak values, at which time the test floor protective circuit breaker opened and the line current dropped to zero. The field amperes and terminal volts returned gradually to zero as the field current decayed. The terminal volts dropped to a low value during the flashover because of line drop of the supply circuit. The voltage to maintain the arc between brush holders was low as compared to what it was to overcome counter electromotive force (emf) just before the flashover.

Fig. 3 is a tracing of an oscillogram produced on a series railway motor by lifting the brushes of one polarity from the commutator. This was done by attaching strong cords to the brush holder fingers and wrapping the brush shunts around the fingers so that both fingers and brushes would be lifted from the commutator. Two assistants were employed to pull on the cords. Since they did not pull exactly at the same time, there were irregular variations of line current for approximately 1/3 second before the flash started. The conditions prior to brush lifting were similar to those for Fig. 1, before interruption, and the general behavior of line current, field current, and terminal voltage during flashover was similar.

Fig. 4 is a tracing of an oscillogram produced on a series railway motor by suddenly throwing a small handful of thin copper slivers on top of the commutator while running. The running conditions again were similar to those of Fig. 1, before interruption, and the general pattern of the current and voltages was also similar.

Fig. 5 is a tracing of an oscillogram of a diesel-electric generator flashover with no overload protection except means for opening the separately excited exciter field at the beginning of the flash, and a circuit breaker in the driving motor circuit. Connections were as in Fig. 6. The flashover was produced by running the generator at full speed, no load, with the main field excited by its exciter to give approximately maximum volts, and then suddenly short-circuiting the main generator terminals through a low resistance. The behavior of the armature current, field current, and terminal volts is typical of a generator flashover produced by a surge of current. The duration of the arcing between brush holders can be readily noted by the rapid irregular variation of these three values. At the time the short circuit was thrown on, the armature current rose very quickly to an excessive value, at which the flashover took place, as shown by the sudden drop and irregular variations of the line current. The determining factor to start the flashover for the set of conditions used was the reaching of a fairly definite value of excessive current. The reason for the sudden drop of line current, is that most of the current between brush holders was flowing through the arc. It should be noted that this is just the opposite to what happens when a series motor flashes over. In that case, the arc current is included in the line current. The arc current in the case of the generator is not measurable because it flows through the armature and then from brush holder to brush holder across the face of the commutator.

The behavior of the generator terminal voltage was to drop suddenly as the current rose. As soon as the flashover became established, the terminal voltage measured was only that across the arc plus the drop in the field coils carrying line current.

The generator main field current shows an induced rise co-ordinate with the initial rise in line current, but decreases in rate of rise as the line current rate of rise decreases. Then when the flashover

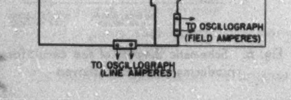

CIRCUIT BREAKER
SHORT-CIRCUITING SWITCH
TO OSCILLOGRAPH (MOTOR VOLTS)
ARMATURE
TO TEST FLOOR SUPPLY CIRCUIT
COMMUTATING FIELD
FIELD SHUNT
MAIN FIELD
TO OSCILLOGRAPH (FIELD AMPERES)
TO OSCILLOGRAPH (LINE AMPERES)

Fig. 2 (left). Schematic diagram of test circuit for producing motor flashovers

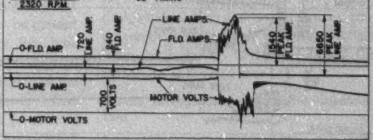


Fig. 3 (right). Tracing of oscillogram of motor flashover produced by brush-lifting

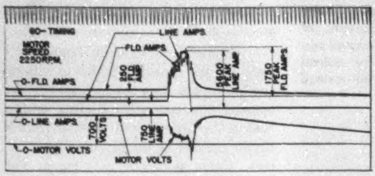


Fig. 4 (left). Tracing of oscillogram of motor flashover produced by copper slivers thrown on commutator

starts, the field current rises very suddenly again to a much greater peak because of excessive armature current, and then oscillates while gradually decaying to the point where the arc goes out. If the peak value of induced field current can be taken as a measure of the rise of the armature current after the flash starts, as it may be before it starts, the value of the arc current is several times the indicated line current at the start of the flashover. This accounts for the potential destructiveness of such an arc if allowed to continue.

As will be seen from the oscillogram, the arcing lasted 22 cycles, or 0.37 second, which is a long time for a destructive arc to continue. In this case the commutator bar ends were badly beaded, the string band blackened with smoke, and brush holder castings pitted. The speed decrease during the flashover was practically negligible because of the large rotating mass of the driving motor.

These typical oscillograms of motor and generator flashovers produced by applying very severe conditions on the test floor give a picture of how the current and voltage values vary before, during, and after flashover. The exceedingly short time for the flashover to start and extremely high values of current obtained are the outstanding features.

Operating Causes of Motor Flashovers

There are many reported causes of motor flashovers in service, although in many cases more causes than one combine to produce the result. These causes can be classified under four general headings: brush jumping, commutator ring-fire, current surges, and bad commutation.

Brush jumping has long been known to be a cause of railway motor flashovers. Hellmund,[1] in 1935, made a rather thorough study of the theoretical factors involved. Brush jumping was a common cause of flashovers on the older type of streetcar and interurban axle-hung 600-volt railway motors that used series-type armature windings and only two brush arms. Sudden rail impacts often caused all the brushes on one of the two brush arms to leave the commutator, thus drawing an arc between brush and commutator causing a condition such as one of the characteristic conditions preceding a flashover. The prescribed remedy usually has been to use a high brush pressure to keep the brushes on the commutator under severe conditions. Brush pressures as high as 9 pounds per square inch (psi) are used on modern railway motor brush holders. Diesel-electric road locomotive motors have multiple-wound armatures, and therefore require as many brush arms as poles (usually four or six). Therefore, track impacts do not have the same tendency to cause all brushes of one polarity to leave the commutator as on the older two-brush arm motors. That brush jumping is one of the causes of diesel-electric motor flashovers, however, is evidenced by the many reports and observations of repeated flashing on certain sections of rough track, or at crossovers.

The condition in which brushes have worn to the point where there is no longer sufficient finger pressure on them should also come under this general heading. Such brushes can leave the commutator surface and cause flashovers even on the smoothest kind of track. Rough or eccentric commutators will cause brush jumping, especially at high speeds if the commutators have high or low spots around the circumference.

Running a motor above its maximum allowable speed can put strains on the commutator bar assembly that are likely to produce a rough commutator and resultant brush jumping. The maximum allowable speed of the locomotive with a certain gear ratio is always definitely stated by the manufacturer. If this is 65 miles per hour (mph), for example, any considerable amount of running above that speed is likely to initiate trouble. A locomotive that has a 65 mph maximum speed is not a 65-mph locomotive. It should be applied in service where there is a margin between the usual top operating speed and the stated maximum speed.

The following operating causes should be classified under the general head of brush jumping:

Eccentric commutator

High bars

Flat spots on commutator

Short brushes

Broken brushes

Running over rough track or railroad crossings at high speed

Exceeding maximum motor speed

Rough or beaded commutator due to previous flashover

Commutator ring-fire as a cause of flashover includes a number of contributing causes. Even new motors with commutators well polished and almost perfectly concentric will develop ring-fire at operating voltages if they have small metallic particles such as copper dust or copper slivers from the undercutting operations left in the slots between bars. Metallic particles such as small steel cuttings are sometimes blown onto the motor commutators with the ventilating air. These are almost certain to produce ring-fire followed by a flashover. Conducting material such as carbon dust, brake shoe dust, or road dust mixed with oil and water are probably the most common causes of ring-fire because of their general prevalence on railway motor commutators.

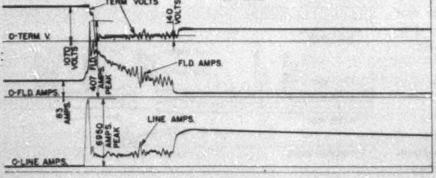


Fig. 5 (left). Tracing of oscillogram of generator flashover produced by short-circuiting through resistance

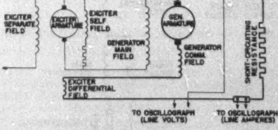

TO OSCILLOGRAPH (GEN. FIELD AMPERES)
SHORT-CIRCUITING SWITCH
EXCITER SEPARATE FIELD
EXCITER ARMATURE
EXCITER SELF FIELD
GENERATOR MAIN FIELD
GEN. ARMATURE
GENERATOR COMM. FIELD
SHORT-CIRCUITING RESISTANCE
EXCITER DIFFERENTIAL FIELD
TO OSCILLOGRAPH (LINE VOLTS)
TO OSCILLOGRAPH (LINE AMPERES)

Fig. 6. Schematic diagram of test circuit for producing generator flashovers

Conducting material on the commutator V-ring extension is conducive to extending ring-fire from brush holder to brush holder.

That wheel-slipping at high speed is one of the causes of motor flashovers is indicated by the increase of the number of flashovers in the autumn on certain lines where wet leaves are likely to be on the track. Certain sections of track known to be slippery for other causes have been reported by trainmen as the locations of repeated flashovers. One such location was found to be where creosoted logs were being dragged over the track.

The following reported causes of flashovers should be classified under the general heading of ring-fire:

Metallic particles in commutator slots
Carbon dust, dirt, oil, or water on commutator
Conducting material on V-ring extension
Conducting material on commutator risers
String band recently painted and placed in operation before drying
Wheel-slipping at high speed
Excessive voltage from generator
Excessive voltage caused by dynamic braking at too high speed
Defective field shunts causing too much shunting of field current

The subject of current surges as a cause has been emphasized in the description of test-floor-produced flashovers. Anything that will cause large and sudden changes in the amperes of the motor circuit can cause arcing at the brushes which may grow into a general flashover. We would include the following reported causes under current surges:

Loose brush shunts which may cause short circuits to ground
Brush shunts touching commutator risers
Too sudden application of dynamic braking at high speed
Improper operation of automatic transition relays or contactors
Short circuits or grounds in wiring or windings of motor
Short circuits or grounds in locomotive wiring external to motors
Main contactors opening and closing under heavy load
Loose connections in main circuits
Rapid change of generator voltage resulting from loose connections in generator field circuit
Rapid surging of load regulator
Rapid surging of engine governor
Reversal of power while locomotive is moving in opposite direction

The degree of sparking at the brushes can be observed easily if the trailing edge of the brushes can be seen. D-c machines that have sparking even to a small degree under steady-state running, usually are criticized. Actually, continuous sparking may be quite severe without in itself producing a flashover. The ultimate result of continued bad commutation is to impair the surface of the commutator, so that one or more of the other causes will produce a flashover more easily. Bad commutation, then, must be listed as one of the general causes of flashovers. The following may be listed as contributing to bad commutation:

Brushes not seated
Lack of brush holder finger pressure
Brushes worn beyond limit of wear
Brushes stuck in brush holders
Improper grade of brush or mixed grades on same commutator
Raw commutator surface
Standstill burns on commutator
Brush holder clearance from commutator not correct
Brush holder moved from correct neutral setting
Running locomotive over maximum speed (overspeeding motors)
Wheel slippage (overspeeding motors)
Excessive dynamic braking current and/or voltage
Field shunts defective so as to cause motors to run in too weak field

Operating Causes of Generator Flashovers

Fundamentally a d-c generator is subject to the same causes of flashover as a d-c motor. On a diesel-electric locomotive, however, the generator does not encounter the same conditions as the motors. The generator does not receive the sharp rail impacts as does the motor, is not subject to extreme overspeeding, and always has the strongest field when its voltage is highest. It may have similar conditions of metallic particles, dirt, oil, or water on the commutator; rough or eccentric commutator; or poor commutation similar to the causes listed for the motors. Unlike the motors, the generator may have overvoltage caused by improper main field setting accompanied by overspeeding caused by improper engine governor setting. Standstill burns and resulting flat spots on the commutator may be caused during engine-starting with a low battery. Voltage and current surges may be initiated in the generator by surging of the load regulator or engine governor. One item of greater importance on generators than on motors is the maintenance of the correct spacing of brush holders from the commutator. Motors, requiring rotation in either direction, use radial brushes, and the reason for keeping the brush holders at the correct distance is to prevent actual rubbing of the commutator if too close, or brush chattering if too far. Since the generator always rotates in one direction, advantage is taken of this fact for using inclined brushes because they give more uniform contact with the commutator. If the brush holder is not at its correct distance from the commutator, the inclined brush is moved circumferentially around the commutator, thus moving the brush off of its correct neutral setting.

A relisting of flashover causes for the generator similar to those for the motor is not necessary since an examination of the motor list will show readily which ones do not apply to generators. This list, however, does not contain the most important cause of generator flashover. By far the most prolific cause of generator flashover is a large and sudden current surge caused by motor flashover. This is confirmed by the fact that few cases of generator flashing are reported where motors have not shown effects of flashing also. When a motor flashover occurs, the result is similar to a near short circuit on the generator. If the current value reached is high enough, the generator will flash over. As soon as this occurs, the voltage at the generator terminals becomes low and the motor flashover quickly dies out. All this occurs in a small fraction of a second. This is the reason many reports read, "Motors lightly flashed, generator heavily flashed." The generator flash will continue and cause severe damage unless means are employed to reduce its voltage and the power input from the engine. Not every motor flashover will cause a generator flashover, but if the sudden surge of current is high enough, it will. It is the writer's belief, based on the examination of many reports, that fully 90 per cent of generator flashovers are initiated by motor flashovers.

Motor and Generator Design Factors Affecting Flashing

Space does not permit complete treatment of this subject. Certain design factors can be controlled in the original design calculation and layout of a d-c generator or motor which will make it less likely to flashover. Principal among the factors to be desired are:

Stable mechanical commutator construction
Low commutator surface speed

Low volts per commutator bar

Low volts per inch from brush to brush

Thick mica separating commutator bars

Low commutation reactance volts, or sparking volts

Brush holder design that will keep brushes always in good contact with commutator

Brush holder arrangement that permits ready escape of ionized gases

Correct brush density

Correct shaping of commutating pole face for sparkless commutation

Wide distance from edge of commutating zone under commutating pole to main pole tips

Low density of commutating pole magnetic circuit at overcurrent conditions

Design of commutating pole magnetic circuit for rapid change of flux

Main pole face shaped to minimize distortion of main pole field form (if pole face windings are not used)

Use of pole face compensating windings to minimize distortion of main pole field form

Arrangement of ventilation to minimize entrance of contaminated ventilating air

All of these factors should receive careful consideration by the designer in the early stages of the design. It is, however, not possible for him unrestrictedly to comply with the requirements of each of the factors listed. In fact, some of these requirements are contradictory. For example, the use of a large number of bars to keep down the volts per bar is in opposition to obtaining low commutation reactance volts. The latter is favored by using fewer armature conductors and hence fewer bars. The use of thicker mica between commutator bars is limited by the thickness of the commutator bar. If the mica thickness is made too great, the mechanical stability of the commutator may be adversely affected. A low commutator surface speed is in opposition to obtaining a large number of bars per pole or low volts per inch from brush to brush. Keeping a wide distance from the edge of the commutating zone to the main pole tip means a lessened percentage of main pole arc. This means a reduction of the rated output that can be obtained from a given size of machine.

Liberality in design of railway-type motors and generators is also definitely limited by the space available. The motor space is limited by the distance between wheel flanges, clearance to the rail, and the distance from centerline of armature to centerline of axle as required for the gear reduction. Generator diameter is limited by aisle space required on each side, and in length by the desire to keep the locomotive length as short as possible. Size of motors and generators is also limited by the desire to keep weight to a minimum.

The designer has to recognize the conflicts among the desired factors, and the space and weight limitations imposed, and to make the optimum compromise of the factors affecting flashing with those necessary to meet a required temperature rating. In a new design it is possible to favor the factors which work toward the elimination of flashing at the expense of those necessary to obtain low-temperature rises. The use of higher temperature insulations has allowed considerable advance in recent designs in this direction. Realizing that the design is a compromise, the best utilization of the factors involved is largely a result of the designer's comprehension of the best relation of all factors, his experience, and his skill in combining them.

In addition to designing for prevention of flashovers, certain features may be used to minimize the effect of flashover if it does occur. One such feature is the use of arcing horns on brush holders and grounding pins adjacent to them. It is much better to ground the arc than to allow it to continue ungrounded. This causes quick action of the ground relay. Also, such an arrangement provides a definite path for the arc to follow and keeps it from hanging on at the commutator bar ends where it is more likely to burn the insulation off the V-ring extension. Fig. 7 shows an arrangement of a grounding pin which has proven effective on diesel-electric generators. The effectiveness of a given arrangement is proven by observing the effect of test floor produced flashovers.

Control Factors Affecting Motor and Generator Flashing

Control design factors affecting flashing, like those for motors and generators, may be either for prevention or for minimizing effects. In the original layout of a control arrangement, attention should be given first to those factors which will tend to prevent motor and generator flashovers. Principal among these are provisions for preventing large and sudden current surges and the selection of motor connections that will allow the generator to operate over its most favorable volt-ampere range. Accepting the fact that flashovers are not completely preventable, the control should also provide means of removing generator voltage and power quickly in case of either a motor or generator flashover.

Consider first the preventative factors. Current surges are mainly associated with the switching operations necessary for motor field shunting, transition from series to parallel motor connections, or changing from motoring to dynamic braking. Current increases resulting from operations of regular field shunting steps are ordinarily not large enough or sudden enough to cause a flashover. Tests on motors have shown that even when running at maximum speed and maximum field shunting, they can be alternately thrown into full field and back into maximum shunting without serious arcing at the brushes. It is usually when some incorrect operation of contactors takes place that flashovers occur in connection with motor field shunting control. A considerable number of flashovers has been reported as a result of improper handling of the transition lever in manual transition. It is undoubtedly for this human-element reason that automatic transition has largely supplanted manual transition. Flashovers also may occur if automatic transition relays and contactors operate improperly. This simply means that these relays and contactors must be designed and built so they will act unerringly.

Sudden and excessive dynamic braking current can cause every motor on a locomotive to flash over. In motoring operations, the motor current is low when the speed is high. When dynamic braking is applied at high speed, the motor becomes a separately excited generator, and only requires sufficient excitation to make it supply both excessive current and voltage. If these excessive values are obtained suddenly, the danger of flashover is greater. The application of dynamic braking can never be taken entirely out of the hands of the engineman, but the

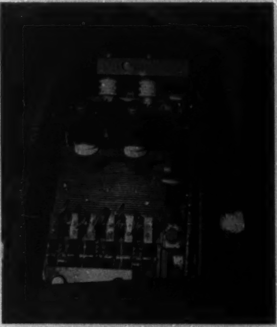


Fig. 7. Brush holder arcing horn and grounding pin arrangement for diesel-electric locomotive generator

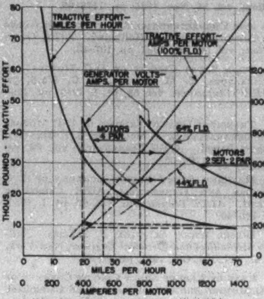


Fig. 8. Curve showing method of plotting locomotive performance for ease of control analysis

control limitation of current and voltage must be definite and positive.

The arrangement of motor connections to the generator and whether or not to use more than one motor connection combination depend on the number of motors used and the amount of power per generator unit. The possible arrangements for a certain locomotive should be studied and compared from the viewpoint of simplicity of control arrangement and the minimizing of conditions which tend toward producing flashovers of motors or generators.

The manner of plotting locomotive performance as shown in Fig. 8, provides a means of showing readily the variation of volts and amperes of motors and generator over the operating range. The changes that occur at field shunting or motor connection change also is readily determined from this form of plot. The usual tractive effort—mph curve is plotted and, in addition generator volts and total tractive effort are plotted against amperes per motor. Tractive effort curves are shown for different values of motor field strengths. The example used in Fig. 8 is for a locomotive with 2,000-horsepower netinput to the generator for traction and using four motors first connected 2-series, 2-parallel, and then 4-parallel. Two steps of motor field shunting are used in the first motor connection, and one in the second. The range of operation of each value is indicated by the heavy lines. This form of plotting locomotive performance may be used for any number of motors and any arrangement of motor connections, and provides a quick method of comparing possible motor connection combinations as regards the factors effecting flashing. The items to be desired are:

1. Operation of motors in as strong a field as possible at high speed. This means better commutator condition over a long period of time resulting from fundamentally better commutation at high speeds. (Sparking volts are proportional to the product of amperes and speed.) The stronger field condition is also fundamentally more stable against flashovers under the conditions that produce brush jumping, ring-fire, or current surges.
2. Operation of generator (and consequently motors) at as low voltage as will obtain the desired performance.

Consider next the control factors that will minimize the effects of flashovers. Much has been done during the last few years to make the control more effective in removing voltage and power from the generator when a flashover occurs. Test-floor-produced flashovers have shown very definitely that the duration of a flashover has a great deal to do with its destructiveness. An uncontrolled generator flashover on a locomotive results in a large amount of power from the engine being dissipated in the arc. This continued for even 1/2 second can be very destructive.

An improved type of ground relay has been developed which is positive not only in detecting a ground anywhere on the power circuit, but which also will detect a flashover on either generator or motor as soon as the arc is grounded. Positive action of the relay is obtained by means of a "biased" alternating current supplied to the ground relay coil circuit through a small transformer.

In addition to the biased ground relay protection, an impulse relay has been developed that operates on the rate of rise of the main generator load current. Since a generator flashover is usually preceded by a very sudden rise of line current, this rise is made to operate the relay. The relay is used to operate the generator main field contactor to insert a relatively high resistance in the main field circuit. This has proved very effective in reducing the duration of a generator flashover.

Fig. 9 is a tracing of an oscillogram of a test-floor-produced flashover on a generator that shows the effect of using this relay. This flashover was produced by short-circuiting the generator at full speed and high voltage in a manner similar to that described for the oscillogram of Fig. 5. The same generator was used in each case. In this case, the short-circuiting resistance was much less than that in Fig. 5, being only 0.019 ohm as compared with 0.106 for Fig. 5. The flashover started approximately 1/60 second after the short-circuiting switch was closed, as shown by sudden drop in line amperes and armature volts. Although not shown in the oscillogram, the impulse relay operated in less than 1/60 second to insert resistance in the generator field. The presence of this relatively high resistance in the main field circuit limited the peak of induced field current to a much lower value than shown in the unprotected flashover of Fig. 5. Also, the field amperes reached a low enough value to cause the arc to go out in approximately 4 cycles of the 60-cycle timing wave as compared to the 22 cycles of Fig. 5. This limitation of the peak generator field current limits the flashover damage, so that repeated flashovers can be produced without damage to the extent that would require removal from service.

What Can Be Done About Flashing?

It probably will continue to be an unsettled question as to whether those who manufacture diesel-electric locomotives and their equipment or those who maintain and operate them can do the most to eliminate flashovers. The view one takes of this usually depends upon which side of the fence he is on. The truth is that both can do more than has been done in the past.

Manufacturers can utilize all of the known factors that will prevent flashing in the fundamental motor, generator, and control designs. They can also incorporate in these designs features which will minimize the effect of flashovers if they do occur. They can arrange the apparatus in the locomotive so that it is accessible for cleaning and adjustment and is unlikely to become dirty and out of adjustment in short periods of time. They can arrange the ventilation air intakes so that clean air is always supplied to rotat-

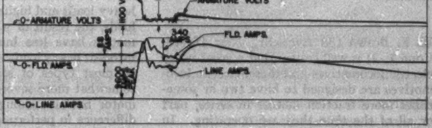


Fig. 9 (right). Tracing of oscillogram of generator flashover produced by short circuit, and limited by impulse relay ing equipment. They can manufacture uniform products so that each machine is as good as any other of the same design. They can supply the operator with thorough instructions as to apparatus construction and the maintenance procedures required. These are admittedly general fundamental statements, but are the goals that should be kept in mind by everyone engaged in the manufacture of diesel-electric locomotives or their electrical equipment.

The operators of these locomotives can reduce the number of flashovers first of all by keeping them cleaner. This is a simple and general statement too, but motors, generators, and control apparatus of the power and voltage required on modern road locomotives will always be more subject to flashovers if wet, oily, and dirty. The extreme contrast between the state of cleanliness of similar equipment in stationary power houses and that on the average diesel-electric locomotive is a distinct shock to the layman who sees the comparison for the first time. Operators can instruct their operating and maintenance personnel more thoroughly in the causes of flashovers and in maintenance to prevent them. They can assign locomotives to service so that there is a margin between the usual maximum operating speed and the stated locomotive maximum speed. They can make needed repairs promptly.

Conclusion

It is unlikely that flashovers on diesel-electric locomotives can be eliminated entirely. Their frequency can be reduced and the amount of resulting damage minimized by a better understanding and use of the causes and remedies. It is hoped that the information given here will contribute toward that desired result.

References

1. Flashing of Railway Motors Caused by Brush Jumping, R. E. Hellmund. *AIEE Transactions*, vol. 54, Nov. 1935, pp. 1178–85.

2. Arc Characteristics Applying to Flashing on Commutators, R. E. Hellmund. *AIEE Transactions*, vol. 56, Jan. 1937, pp. 107–13.

3. Solid Short Circuit of D-C Motors and Generators, T. M. Linville, H. C. Ward, Jr. *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 119–24.

4. Transient Characteristics of D-C Motors and Generators, A. T. McClinton, E. L. Brancato, Robert Panoff. *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 1100–06.

5. Maximum Short-Circuit Current of D-C Motors and Generators, AIEE Committee Report. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 146–49.

6. Diesel Wheel Slippage, F. Thomas. *Railway Mechanical and Electrical Engineer*, New York, N. Y., vol. 125, Oct. 1951, pp. 80–84.

7. Rate of Rise of Short Circuit Current of D-C Motors and Generators, AIEE Committee Report. *AIEE Transactions*, vol. 71, pt. III, Jan. 1952, pp. 314–25.

8. Short-Circuit, Flashing of D-C Machines, A. T. McClinton, J. P. O'Connor. *AIEE Transactions*, vol. 72, pt. III, Feb. 1953, pp. 1–7.

9. Flashover Torque of a D-C Generator, O. C. Coho. *AIEE Transactions*, vol. 72, pt. II, March 1953, pp. 43–48.

10. Report of 1953 Diesel-Electric Committee, Locomotive Officers Maintenance Association. *Railway Locomotives and Cars*, New York, N. Y., Oct. 1953, Topic "A"—Flashovers, pp. 76–7.

◆

Discussion

J. K. Stotz (Fairbanks, Morse and Company, Beloit, Wis.): I agree with Mr. Atwell that flashing can never be entirely eliminated and that better maintenance will reduce its frequency and possibly the resultant damage. However, the basic equipment must be so designed as to commutate satisfactorily over the entire operating range, and that means up to the maximum safe speed and to current values corresponding to 25 to 30 per cent adhesion at least, because in spite of the electrical manufacturers' desires, locomotives are going to be operated over this range and the railroads will expect and should get satisfactory performance.

Motors and generators designed as conservatively as possible from the electrical standpoint will still flash under some conditions even with the best maintenance. Their mechanical design should be such that the arcs produced by the flashover will do a minimum of damage. This means the elimination of pockets confining ionized gases over insulated surfaces, the complete insulation of field coils to prevent arc spatter short-circuiting adjacent turns, which will thus cause bad commutation after a flashover and the use of brush holder and V-ring insulation that resists carbonizing when hit by a flash. The protective devices must also be sensitive and fast-acting further to minimize the damage caused by flashovers.

H. F. Brown (38 Avon St., New Haven, Conn.): It is well known that a great many electric locomotives and diesel-electric locomotives are designed to have two or sometimes more traction motors in series, part or all of the time they are operating. In most of these cases, the motors are connected to different axles, so that when one pair of driving wheels slips, the voltage is unequally divided across the two motors in series. The author does not specifically include wheel-slipping with motors connected in series in his category of causes of motor flashovers.

Has it been determined from operating experience or from laboratory tests that flashovers on traction motors occur more frequently when the motors are connected in series than when they are in parallel? Probably the effect of such slipping in causing flashovers is approximately the same; but it would appear however that the tendency to slip would be less with the motors in parallel.

C. F. Jenkins (Westinghouse Electric Corporation, E. Pittsburgh, Pa.): The author's comprehensive outline of the causes of flashing on d-c motors makes a good background against which to inquire as to why on diesel-electric locomotives motor flashing is worse than on 3,000-volt trolley locomotives. This occurs in spite of the fact that the available power is lower on the diesel. Also, the possibilities of voltage surges and circuit interruptions are practically nil, and the maximum peak volts per bar on the commutator are about 20 per cent lower.

One point of difference is the constant horsepower characteristic of the diesel locomotive. This results in lower speeds at heavy loads and higher speeds at the lighter loads, the result of which is that the diesel motors have less burning of the commutators at heavy loads and more at light loads. In most types of service this may mean somewhat more severe service on the diesel motor but not enough to account for the difference in performance.

Another point of difference is the type of protection provided. The trolley motors are protected by means of high-speed overload trips which disconnect the motor from the line when flashover occurs. This is very effective in limiting the damage in any one flashover. However, as the author points out, the diesel motor is protected in somewhat the same manner by reason of the fact that the motor flashover usually causes the generator to flash, thus short-circuiting its terminals, reducing the motor voltage, and extinguishing the arc.

As far as the motors are concerned, this leaves the accumulation of oil with its propensity for collecting carbon dust and other forms of dirt probably the greatest single factor contributing to the greater tendency toward flashing of the diesel motors. This is understandable when it is considered that conducting particles between bars can cause flashover of themselves, and therefore it is very probable that they will contribute to flashover tendency during bad commutation caused by any other factor such as brush jumping, current surges, or wheel slip.

This is a cause which can be eliminated in part by the operators in keeping the equipment clean and free from excess oil, but much more can be accomplished by the locomotive designer by arranging his equipment so that none of the ventilating air is drawn from the diesel engine compartment or other contaminated source. As a matter of fact this is now being done by one European locomotive builder.

C. A. Atwell: I agree with Mr. Stotz that the motors and generators must be designed to commutate satisfactorily from 25 to 30 per cent adhesion (when geared for freight service) up to the maximum specified speed. However, as on any kind of transportation drive equipment, a lot of running at or close to the maximum specified speed means less life and more maintenance. Railroad operators can obtain more mileage from their equipment by applying locomotives with a motor gear ratio so that there is some margin (say 10 or even 15 per cent) between the maximum sustained speed and the manufacturer's specified maximum speed.

In answer to Mr. Brown's question, the reason a pair of wheels geared to an axle slips, is that the tractive effort is greater than the coefficient of friction will stand. This is independent of how the motors are connected. When one of two motors connected in series does start to slip, its higher speed causes it to generate more counter-electromotive force and therefore to take more than its half of the total voltage across the two motors. This extra voltage naturally tends toward flashover if the slipping speed is high enough. If this motor does not flash over, the other one has less than its share of voltage and is less likely to flash over. As long as the second motor does not flash over, it will act as a limitation to the short-circuiting effect on the generator of a flashover on the first motor. If both motors in series happen to slip to high speed at the same time, they both may flash over and provide a near short-circuit on the generator.

Another point to note is that when a motor slips to high speed, its increased counter-electromotive force reduces the current it takes from the generator. If a second motor is in series with it, it will also have less current. At reduced generator load current the voltage of the generator is higher because of its differential characteristic. This affects any parallel motor circuit and may cause motors in it to tend to slip. However, the more parallel circuits, the less of this effect there is.

There is some advantage as regards flashover in having two motors in series and some in having all motors in parallel. These advantages are not easily evaluated because of the many different combinations of slipping axles that occur in service and the difficulties in observing flashovers resulting from wheel slipping. Many wheel slips have been observed in service but observed flashovers resulting from them have been few. All of our efforts to cause flashovers on locomotives by intentionally slipping wheels at high speed, or by simulated wheel-slipping in the laboratory, failed to produce a flashover.

Mr. Jenkins' comments on traction motors on diesel-electric locomotives as compared to those on trolley locomotives are well taken, as are his comments stressing the effect of the presence of oil and dirt on motor commutators as a cause of motor flashovers. Maintenance that keeps motor commutators free from oil, water, and dirt is probably the most fertile field for prevention of flashovers on diesel-electric locomotives.

Co-ordinated Fuse Protection for Low-Voltage Distribution Systems in Industrial Plants

JOHN C. LEBENS
MEMBER AIEE

UNTIL the advent of the dual-element fuse, co-ordinated fuse protection for low-voltage distribution systems was virtually impossible because the time-current characteristic of the nonrenewable and renewable fuse did not match the safe time-current characteristic of the circuit. Even short-circuit protection on a-c distribution systems was doubtful because of the lack of data on the interrupting capacity of nonrenewable and renewable fuses. The greatly enhanced time-lag characteristic of the dual-element fuse makes it possible to protect the low-voltage system against dangerous overloads and its interrupting capacity assures satisfactory performance under fault conditions. Since dual-element fuses only are made in National Electrical Code sizes, the Hi-Cap fuse was designed in ratings up to 5,000 amperes to fill co-ordinated fuse protection needs.

Until recently, power fuses which were rated at 600 volts or less, and commonly referred to as low voltage fuses, offered little more than short-circuit protection to the distribution system. The nonrenewable cartridge fuse is the oldest and most common type. Fundamentally it consists of a zinc link surrounded by an arc-quenching filler. Such fuses have very low thermal capacity so that once the current exceeds their rating they blow rapidly.

The renewable fuse represents an improvement over the nonrenewable fuse in that, with it, the cost of fuse replacement is reduced even though the original cost is greater. The arc-quenching filler is eliminated by building the case stronger, thereby simplifying the renewal of the fusible element or link. Even more important, the thermal capacity of the fuse link is made as large as possible so that the fuse will hold harmless transient currents without blowing. However, there are physical limits to the amount the mass of the renewal link can be increased without developing mechanical or electrical troubles so that the amount of time lag which can be built into the renewable fuse is limited.

For this reason the size of the nonrenewable or renewable fuse selected for the protection of distribution circuits feeding motor loads is equal to three to four times that of the load to prevent the fuse from blowing under normal, harmless transient conditions produced by the starting of motors, switching of capacitors, or the like. Hence, the nonrenewable or renewable fuse is capable of giving only short-circuit protection, and overload protection requires an additional device.

The dual-element fuse overcomes the limitations of the nonrenewable and renewable fuses. As shown in Fig. 1, it consists of a short-circuiting strip surrounded by arc-quenching filler in each end of the fuse with a thermal overload device in the center. Current flowing through the fuse generates heat in the short-circuiting strips. At the lower overloads, where there is time for thermal conduction, heat is conducted into the thermal overload device in the center, raising its temperature to 280 degrees Fahrenheit (F), at which point the fusible alloy melts, permitting the interrupter to move

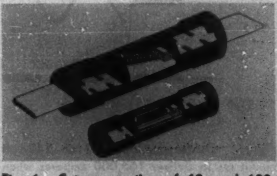


Fig. 1. Cutaway section of 60- and 100-ampere 250-volt dual-element fuse

---

Paper 54-9, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted September 22, 1953; made available for printing October 27, 1953.

JOHN C. LEBENS is with the Bussmann Manufacturing Company, St. Louis, Mo.

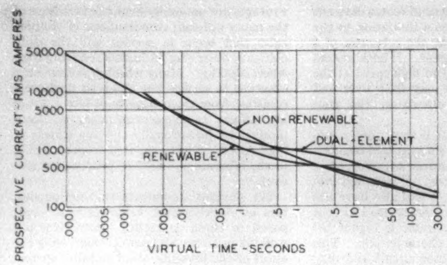


Fig. 2 (left). Time-current characteristic of 100-ampere 600-volt nonrenewable, renewable, and dual-element fuses

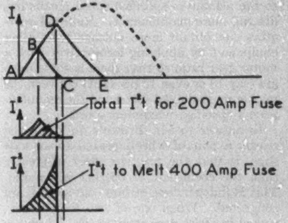


Fig. 4. Co-ordination between a 200- and a 400-ampere fuse when the current flowing is being limited by the smaller fuse

mechanically, under spring tension, opening the circuit. At the higher overloads, where there is not time for thermal conduction, the heat generated in the reduced sections of the short-circuiting strips raises their temperature to the melting point causing them to operate like ordinary fast-acting fuses.

By combining a thermal cutout with fast-acting fuse links in the same fuse case, sufficient time lag can be built into the device to eliminate needless blowing on harmless overloads and still obtain fast action under short-circuit conditions as shown in Fig. 2. The comparison between the dual-element fuse, the ordinary nonrenewable fuse, and the renewable fuse clearly shows the effect of the thermal cutout of the dual-element fuse on its time-current characteristic.

All three devices have the same current rating. All three are capable of carrying 100 amperes indefinitely and will open at 135 amperes in approximately the same time. However, from this common starting point the characteristics of the three devices diverge as shown. The ordinary nonrenewable fuse, having relatively little thermal capacity and only one mode of operation, has a smooth time-current characteristic without any points of inflection. It has little time lag at the lower overloads so that it will open on harmless transients. Even so it is slower than the dual-element fuse under short-circuit conditions.

The renewable fuse, with its increased thermal capacity has longer time lag than the nonrenewable fuse at the lower overloads but the effect of the increased mass is lost above 300 per cent load. Hence, it represents little improvement in the protection of motor circuits. Since the mode of operation changes with the overload, the curve contains a point of inflection and its opening time under short-circuit conditions is less than that of the nonrenewable fuse. The dual-element fuse has even more time lag at the useful loads but at fault currents it drops to an amazingly short time.

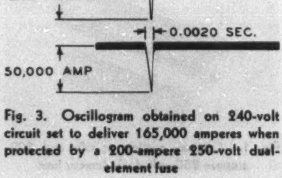


Fig. 3. Oscillogram obtained on 240-volt circuit set to deliver 165,000 amperes when protected by a 200-ampere 250-volt dual-element fuse

The greatly enhanced time-lag characteristics of the dual-element fuse makes it possible to match the branch circuit fuse to the actual load even if it is a motor load with poor starting characteristics. This greatly simplifies the co-ordination problem because, as the fuse size increases in going from the branch circuit back through the distribution circuits to the main entrance equipment, a greater spread between fuse sizes is possible. This reduces the possibility of the larger, feeder fuse blowing instead of the branch circuit fuse under fault conditions.

Even more important, the dual-element fuse has simplified the problem of furnishing co-ordinated fuse protection for low-voltage distribution systems because its time-current characteristic is not affected by service. The maximum operating temperature of the dual-element fuse is limited by the melting temperature of the fusible alloy used in the thermal cutout which is 280 F. At this temperature the component parts of the fuse are not subject to oxidation or erosion so that the fuse can carry full load indefinitely without damage. For the same reason any overload which is not continued long enough to cause the thermal cutout to operate does not damage the fuse or affect its time-current characteristic. Hence it truly differentiates between safe and harmful overloads. In the final analysis the demarcation between satisfactory and faulty operation is a matter of time at any particular value of current. If the current is cut off before damaging conditions are reached, the electrical circuit is not being operated at its full capacity, whereas if the current is allowed to continue too long, the circuit and the associated equipment are destroyed. Hence, for complete protection, the time-current characteristic of the protective device must be matched to the safe time-current characteristic of the circuit and this time-current characteristic must remain the same throughout the life of the fuse.

The co-ordination of the dual-element fuses in the overload range is a simple matter because the sizes can be selected by referring to time-current characteristics, it can be checked by actual laboratory test if desired, and a relatively large separation between opening times is possible because the arcing time represents only a small percentage of the total time. However, under fault conditions where the dual-element fuse blows in its short-circuiting strips and the time of opening is extremely short, the co-ordination problem requires more careful analysis.

For adequate protection the fuse must have sufficient interrupting capacity to handle safely the fault currents to which it may be exposed and its time-current characteristic under these conditions must be established so that the protective device nearest the point of trouble will open, thus minimizing the disturbance and damage and preventing unnecessary outages to the rest of the system.[1] Until recently the technical data available on low voltage fuses were insufficient to establish either requirement.

The only short-circuit tests conducted on low-voltage fuses by unbiased laboratories were those of the Underwriters' Laboratories on direct current with 10,000 amperes available. Even though field experience with fuses has been unusually successful, no basic information

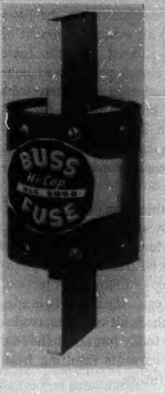


Fig. 5. Hi-Cap fuses available in ratings up to 5,000 amperes at 600 volts or less

as to the performance of the fuse on heavy alternating fault currents was available. Hence neither the interrupting capacity nor the time-current characteristic at these currents was established.

To obtain these data tests witnessed by the Electrical Testing Laboratories of New York were conducted on 250- and 600-volt Fusetron dual-element fuses on circuits capable of delivering peak currents as high as 235,000 amperes at 240 and 545 volts. These tests established that the most severe condition occurred on a single-phase short-circuit with only one fuse in series with the short circuit. Under these conditions all the Fusetron fuses tested cleared the circuit safely and current limitation occurred when the available current was approximately equal to 40 times the rating of the fuse. This current limitation was obtained because the fuse opened the circuit in less than 1/4 cycle, before the current could build up to the available amperes of the circuit.

A typical oscillogram is shown in Fig. 3. This picture of the 200-ampere 250-volt dual-element fuse on a 240-volt circuit set to deliver 165,000 peak amperes shows an instantaneous peak current of 50,000 amperes, a total clearing time, including the melting and arcing times, of 0.0020 seconds and a peak transient voltage of 840 volts. These tests prove that the Fusetron dual-element fuse has an interrupting rating in excess of 100,000 rms amperes but, without interpretation, it is difficult to use in a co-ordination study.

To illustrate, consider the condition resulting when a 200- and a 400-ampere dual-element fuse are in series with a calculated fault current of 100,000 amperes. Assume that the fault develops at current zero as shown in Fig. 4. The current builds up along the sinusoidal curve to some point such as B if only the 200-ampere fuse in the circuit and to some point such as D if only the 400-ampere fuse is present. In either case when the dual-element fuse blows the current decays along a line such as BC or DE.

Co-ordination is obtained if the 200-ampere fuse blows to clear the fault before sufficient current flows to melt the 400-ampere fuse. If the instantaneous current flowing is squared, the energy developed while the 200-ampere fuse is blowing is equal to the area under the curve as shown. Since this is less than the amount of energy required to melt the 400-ampere fuse, the fuse is not damaged and proper co-ordination is obtained.

Rather than force the circuit designer to perform this calculation for every available fault current and combination of fuses, a better plan is to develop time-current characteristics over the entire range based upon the current which would flow if not limited by the fuse. The International Electrotechnical Commission has adopted the term "prospective current" to describe this current.[2]

The prospective current is no different from the symmetrical rms "available amperes" now used. In other words, it is the current which would flow if it were not limited by the protective device. It is controlled entirely by the circuit constants.

The Commission has also adopted the term "virtual time" to describe the time that the prospective current must flow to produce the same heating effect in the circuit as that produced by the actual current flowing while the fuse is blowing. By means of the prospective current and the virtual time the complete time current characteristic can be established from the extremely long blows to instantaneous blows under the heaviest fault conditions, as shown by the dual element curve of Fig. 2.

For co-ordination studies a virtual melting time and a virtual arcing time must be used. The virtual melting time is the time that the prospective current must flow to produce the same heating effect as that produced by the actual current flowing while the fuse is melting. The actual current during this period is represented by the curve from A to B in Fig. 4. In like manner the virtual arcing time is the time that the prospective current must flow to produce the same heating effect as that produced by the actual current flowing while the fuse is clearing the circuit. The actual current during this period is represented by the curve from B to C in Fig. 4.

If at any given value of prospective current the virtual melting plus arcing time of the smaller fuse is less than the virtual melting time of the larger one, proper co-ordination is obtained.

The dual-element fuse is made only in the sizes established by the National Electrical Code so that it is limited to ratings of 600 amperes and less. But in many industrial distribution systems fuse ratings as large as 5,000 amperes are required to complete the co-ordinated protection picture. For this reason the Hi-Cap fuse, as shown in Fig. 5, was developed. It is designed to be bolted directly to the bus bar rather than installed in fuse clips.

The Hi-Cap fuse uses fine silver links embedded in chemically inert arc-quench-

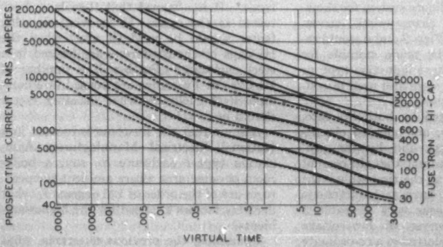


Fig. 6 (left). Time-current characteristics of Fusetron and Hi-Cap fuses. Dotted lines, 250 volts; solid lines, 600 volts

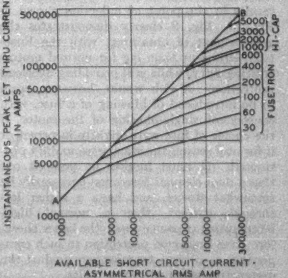


Fig. 7. Current limiting effect of Fusetron and Hi-Cap fuses

ing filler. The terminals are a high copper alloy, silver-plated, sealed to the tube with *O* rings and held in position with screws. A glass-cloth laminated-base melamine-impregnated tube is used because of its excellent arc-quenching properties, high temperature resistance, and dimensional stability.

The fuses are rated at 600 volts or less in sizes up to 5,000 amperes. The time-current characteristic is established so that the Hi-Cap fuses co-ordinate with the dual-element fuses even though the Hi-Cap fuse has less time-lag at the lower overloads. Since the load on the Hi-Cap fuse in any distribution system is more diversified than the load on the dual-element fuses used in the sub-feeders and branch circuits the need for the longer time lag at the lower overloads is not required.

The time-current characteristics of the Fusetron dual-element fuses and the Hi-Cap fuses in ratings from 30 to 5,000 amperes (Fig. 6) show the co-ordination obtainable with these devices even though the time lag of the Hi-Cap fuse is less than that of the dual-element fuses at the lower overloads. Since the Hi-Cap fuse is intended for short-circuit protection only, the overload range is of little importance in its application. However, the watt loss in the fuse while carrying full-load current is extremely important because, if care if not exercised in the design, excessive temperatures in the fuse result.

The ideal would be realized if the Hi-Cap fuse could be replaced with a section of bus bar and if the temperature rise of the bus bar could be no lower than that of the fuse when carrying the same current. Actually this condition is approached with the Hi-Cap fuse. A 2,000-ampere Hi-Cap fuse when carrying rated load showed a temperature rise on the terminal of 60 degrees centigrade (C) whereas a 2,000-ampere bus bar, replacing the Hi-Cap fuse, showed a rise of 32 C when carrying the same current. Hence, the fuse itself contributed enough heat to produce a rise of only 28 C.

Interrupting-capacity tests witnessed by the Electrical Testing Laboratories show that the Hi-Cap fuses also have an interrupting rating in excess of 100,000 amperes. Tests conducted on the fuses at 535 volts with an available current of 212,000 amperes showed that the fuses cleared the short circuit perfectly without any external disturbance. In these tests the Hi-Cap fuse limited the let-through current as shown in Fig. 7. As would be expected, the current limiting effect of the fuse is more pronounced in the smaller sizes.

Hence, the combination of the Fusetron dual-element fuse and the Hi-Cap fuse, when applied to low-voltage distribution systems in industrial plants, makes possible completely co-ordinated protection. The interrupting capacity not only is adequate to handle the fault currents available on the system today but also to handle the short-circuit currents resulting from expansion in the future.

References

1. ELECTRIC POWER DISTRIBUTION FOR INDUSTRIAL PLANTS. *AIEE Special Publication S-3.*

2. THE APPLICATION AND STANDARDIZATION OF HIGH RUPTURING CAPACITY CURRENT-LIMITING FUSES, J. W. Gibson. *AIEE Transactions,* vol. 72, pt. II, May 1953, pp. 126–32.

◆

Discussion

G. W. Gibson (Electric Transmission, Ltd., Etruria, Stoke-on-Trent, Staffordshire, England): The paper well demonstrates how improvements in fuses and increased knowledge of their behavior widens their sphere of application. A very interesting example is the author's scheme of co-ordinated protection by Fusetron and Hi-Cap. Is there now a tendency in the United States to eliminate circuit breakers from some low-voltage systems? This has long been done in Britain where it seems appropriate, but British practice has always been to avoid large concentrations of low-voltage power, by means of sectionalizing, for the purpose of limiting fault currents.

Referring to the renewable fuse, what is the change in the mode of operation which results in the point of inflection shown in Fig. 2? Fig. 2 clearly demonstrates the achievement of obtaining, with the Fusetron, sufficient time lag for normal motor-starting duty, while still providing overload protection. A risk would, however, seem to be introduced of blowing of a fuse, with consequent single-phasing of the motor, if the choice of fuse were slightly incorrect or if for any reason the start were unduly protracted. It must further be pointed out that a dual-element fuse must necessarily be somewhat slower, and have a higher let-through current, than a normal high-interrupting capacity fuse; the latter therefore gives improved protection to such components as current transformers against the destructive effect of heavy faults.

The author's remarks regarding the temperature rise of the 2,000-ampere Hi-Cap fuse are interesting. With fuses of this and higher current ratings, the problem of heat evolution, and so of possible contact deterioration, may become important. I therefore think the author is wise in using silvered contacts. The higher the current rating of a fuse, the greater is the proportion of the heat generated by it which is dissipated by transmission along the attached conductors, rather than from the fuse barrel. It would therefore be interesting if Mr. Lebens would give an approximate value of this percentage for Hi-Cap fuses of the higher current ratings.

L. F. Ferri (Commonwealth Associates Inc., Jackson, Mich.): The paper is an important contribution to the knowledge of high-current rated low-voltage fuses. It is well to know that the problem of selectivity so well appreciated by the relay engineers is also recognized by the manufacturers of low-voltage fuses. It would however be more gratifying if a common base could be agreed upon by the engineers concerned with fuse protection. It would then be possible to adopt a similar co-ordinate system for showing the time-current characteristic curves. The curves shown in Figs. 2 and 6 would require redrawing to be made suitable for selectivity studies where induction-type or other protective relays are also used in conjunction with these fuses.

It has been accepted by both the fuse, (over 600-volt) relay, and molded circuit-breaker manufacturers, and by the protection engineers in general, that the current scale is the abscissa and the time scale is the ordinate of the time current characteristic curve. It would appear very desirable if the characteristic curves of low-voltage power fuses could be plotted in a co-ordinate system which is familiar to all the relay engineers.

William J. Conlon (Guy B. Panero Engineers, New York, N. Y.): The paper indicates that it is now feasible to design a fully co-ordinated system of protection by fusing with Hi-Cap fuses and Fusetron element. The paper does not present any data on the co-ordination of the overload characteristics inherent in the Fusetron with the overload element in magnetic motor controllers and manual motor controllers. It would be interesting to have data of these two features. For example, 440-volt squirrel-cage induction motor should be protected by a 30-ampere Fusetron. Assume the motor drives a compressor controlled by means of a magnetic controller which in turn answers to the demands of a thermostat. In the case of a 50-per-cent overload which device will open first? Also consider other percentages. Curves would be helpful.

No statement has been made concerning the method of rating Hi-Cap fuses or Fusetrons. It is assumed that they have been tested in the open. If so, what derating factor should be used (if any) when the Hi-Caps or the Fusetrons are housed in a steel enclosure? Further if a number of Hi-Caps or Fusetrons are housed in a common steel enclosure what derating factor should be used?

Assume that the Fusetron is located in a National Electrical Manufacturers Association type-*1* enclosure in, say, a boiler room or some area where ambient temperatures are of the order of 125 degrees F, what derating factors are used for the higher ambients?

Combining the previous questions, what

derating factors are used in cases where a number of Fusetrons are housed in a common sheet steel enclosure in an area where ambient temperatures are in excess of those normally not in common use?

Assume that an existing building is equipped with thermal-type circuit breakers in panel boxes for branch circuit protection of lighting circuits. The building will be expanded and a new feeder and sub-feeder will be installed. Is it possible to co-ordinate Hi-Caps or Fusetrons with the existing thermal breakers?

R. H. Kaufmann (General Electric Company, Schenectady, N. Y.): The basic objective of this interesting paper was the achievement of a longer time delay in the moderate overcurrent range with the least possible increase in time delay in the high overcurrent range. The availability of a reliable dual-element fuse with this enhanced time-current curve would extend the application possibilities.

The achievement of this goal requires that unique design and application problems be solved. The allocation of only a portion of the fuse cartridge to the high-current element increases the internal stress under short-circuit interruption, and requires a corresponding increase in strength to avoid explosion. The longer time-delay imposes greater stress on devices which rely on the fuse for protection.

In connection with the short-circuit behavior of this dual-element fuse, the author reports tests which were witnessed by the Electrical Testing Laboratories. Can he tell us if the test performance was satisfactory? In testing individual 600-volt fuses in a 545-volt 60-cycle a-c circuit with an available peak current of 235,000 amperes, was a pressure explosion of the fuse barrel observed? Would such pressure explosion be expected if tested at lower available fault currents? Is performance rated as satisfactory if the current is interrupted, even though the fuse explodes?

A circuit in which the fuses were tested is described as one capable of delivering peak currents as high as 235,000 amperes. If interpreted literally, this defines a circuit capable of delivering an rms symmetrical current of about 90,000 amperes or one with an interrupting capacity (IC) in accordance with AIEE calculating procedure for low-voltage circuit breakers of about 115,000 amperes. Since most published data relative to low-voltage system available short-circuit currents are expressed in the last mentioned terms, it would be proper to express the available capacity of a fuse test circuit in like terms. One wonders whether all reference to test circuit capacities are expressed in terms of maximum crest amperes.

In a concluding paragraph Mr. Lebens states that tests performed show that the fuses have a rating in excess of 100,000 amperes. A rating is an assignment by a manufacturer certifying that the device will meet stated performance standards. Tests serve to prove the rating but do not in themselves impart a rating to the device tested.

The title of the paper is "Co-ordinated Fuse Protection for Low-voltage Distribution Systems in Industrial Plants" and the concluding paragraph states that the combination of the Fusetron dual-element fuse and the Hi-Cap fuse make possible completely co-ordinated protection. Yet a review of the paper discloses that the only co-ordination property discussed is that pertaining to selective operation of fuses in series; and that only partially. The paper states that selective operation will be achieved if the I^2t total let-through of the far fuse does not exceed the I^2t melting value of the near fuse. This is roughly true but needs some expansion. A fuse link can be damaged or partly melted without opening. Instead of the melting characteristic, reference should be made to the damage boundary which will be somewhat beneath the melting characteristic. Since individual fuses are subject to manufacturing tolerances, assured selective operation requires comparison of the maximum total clearing characteristic of the far fuse with the minimum damage characteristic of the near fuse. Fig. 6 presents a time current curve without defining whether it represents the damage boundary, minimum melting or maximum total clearing. Would Mr. Lebens identify this curve and state where the other curve or curves necessary for establishing selective operation can be obtained?

In passing, it will be well to note that the horizontal available current scale in Fig. 7 is plotted in terms of the total asymmetrical rms current in one phase which is not the short-circuit IC value so widely available in handbooks and tables. I suspect that the available current values in Fig. 7 represent 1.4 times the rms symmetrical current. Thus, values of available IC as tabulated for low-voltage circuit-breaker application in accordance with AIEE standards should first be multiplied by the ratio 1.4/1.25 before entering the curves of Fig. 7. A confirmation or correction of this procedure by the author would be helpful.

There are other co-ordination characteristics which must be dealt with in acquiring co-ordinated protection. First is the item of a co-ordinated switch unit. The fuse being an expendable interrupter must be replaced after each operation. To perform this operation safely demands that the fuse assembly be de-energized. Fully to realize the need for this function one need only recall the records in *Electrical Work Injuries in California Industries (1940–1951)* by E. E. Carlton, compiled by the State of California and presented in an unpublished AIEE paper in 1952. During a 10-year interval, electrical work injuries in California industries accounted for 4,851 personal injuries and 380 deaths. Mr. Carlton reports that about 80 per cent of the injuries, or around 3,900, occurred on low-voltage circuits (600 volts and less). He further reports that about one out of every five injuries was the result of failure to de-energize equipment before maintenance or repairs were begun.

That careful protection co-ordination between switch and fuse is most important is also indicated by the same set of accident records. Mr. Carlton reports that about one out of every 6 injuries is caused by switch or switch-fuse combination failure, most of which were actual explosions. In the light of such performance records, installations in major industrial distribution circuits should be made only after the most thorough and exacting co-ordination studies. The paper fails even to mention the problem, to say nothing of indicating the procedure to be followed to insure protective co-ordination. Perhaps it was felt that insufficient data were available on which to base a proposal. Mr. Lebens should be encouraged to proceed in establishing safe co-ordination standards.

One of the major application fields for the low-voltage current-limiting fuse is the protection of other devices connected in the electrical circuits which cannot withstand unrestricted short-circuit current flow. The let-through current of the fuse becomes greater, the larger the ampere rating. Thus a particular circuit device or element could safely withstand the current let-through of some particular ampere-rated Fusetron dual-element fuse but would be damaged or destroyed by the let-through of one of higher ampere rating.

The let-through current of the fuse could be defined by appropriate total clearing curves which would resemble those of Fig. 6 and probably could be made available by the fuse manufacturer. The difficult part of the co-ordination problem is in defining the withstand ability of the devices to be protected. Such characteristics can be established by full-scale testing but this is expensive and often involves an indecisive situation as to who is certifying and accepting the responsibility for the withstand ability of devices to be protected. No mention of this co-ordination problem is made in the paper. Some word of guidance for a prospective user indicating the way or ways in which adequate protective co-ordination can be assured would be very welcome.

The subject of short-circuit protection co-ordination should not be treated lightly. Improper protection co-ordination can lead to costly equipment damage and lost production time and even personal injury or loss of life. If you have any doubt about the truth of this statement, study carefully the California industrial electrical accident reports, particularly individual annual reports which contain detail classified listing of particular types.

Let me hasten to reassure everyone that electrical systems are not inherently dangerous or hazardous to property and life. Misbehaviors are invited by applying equipment which has not been certified as capable of meeting the maximum service duty. By application of a sound intelligent engineering approach to the protection coordination problem the full potential capabilities of protective devices can be utilized safely.

It is gratifying to see the low-voltage fuse industry becoming cognizant of the importance of adequate short-circuit performance. Their continued interest and study when reflected in proper application procedures will lead to greater safety in industrial electrical systems.

John C. Lebens: The author is appreciative of the excellent discussions presented. Unfortunately space does not permit a complete discussion of all the questions raised.

Mr. Gibson well expresses the English viewpoint and asks if there is now a tendency in the United States to eliminate circuit breakers from low-voltage systems as is done in England. Actually low-voltage power systems in the United States have been installed using fuses alone, circuit-breakers alone, and combinations of fuses and circuit breakers. The trend toward larger capacity circuits with higher available short-circuit currents has increased the use of fuses to protect inadequate circuit breakers as well as the circuit itself.

Mr. Gibson asks what the change is in the mode of operation of the renewable fuse which results in the point of inflection shown in Fig. 2. The type of renewable fuse referred to consists of a fuse link with three reduced sections or weak spots separated by wide sections whose mass is increased by lag plates welded to them. The size of the weak spot in the center of the link is larger than the two weak spots near the fuse terminals.

On loads up to 300 per cent where there is time for thermal conduction, the end weak spots are cooled so that the blow occurs in the center weak spot and the time lag is dependent upon the thermal capacity of the wide sections and lag plates. At loads greater than 300 per cent there is insufficient time for thermal conduction so that the blow shifts from the larger center weak spot to the smaller ones at the terminals, hence, the point of inflection in the time-current characteristic.

With reference to the danger of single-phasing a motor protected with dual-element fuses whose ratings have been selected substantially equal to the motor rating, when the motor starts slowly the single-phase condition does not develop because all the dual-element fuses have been heated equally by the starting current so that, if one blows, the others also are ready to open. If the fuses do not blow simultaneously the single-phase condition produced by one fuse opening causes the current to increase in the other two lines, blowing the remaining dual-element fuses protecting the motor.

Mr. Gibson's point is well taken that, for any given rating, the let-through current under heavy short-circuit conditions is greater for the dual-element fuse than for a normal high-interrupting capacity fuse as known in England. However, it is unreasonable to compare the dual-element fuse with the normal high-interrupting capacity fuse, ampere for ampere, because they are not applied in the same manner.

In the protection of industrial circuits feeding motors and similar loads the dual-element fuse is selected so that its rating is substantially equal to the motor name-plate rating, whereas the normal high-interrupting capacity fuse rating is selected equal to three or four times the motor rating. Hence, in any comparison of let-through currents, the dual-element fuse must be compared with a normal high-interupting capacity fuse three or four times its size. In such a comparison the let-through currents are found to be substantially the same.

I heartily agree with Mr. Ferris' comments that anything that can be done to standardize the method of representing the time-current characteristics of the various circuit components is highly desirable. As long as times greater than 1 cycle are considered, the representation of the relay, circuit-breaker and high-voltage fuse suppliers is satisfactory. However, the system breaks down under 1 cycle and becomes meaningless.

Since the International Electrotechnical Commission already had run into the trouble and solved it, we adopted their representation in a step toward international standardization. Knowing the natural inertia of the American engineer, we contemplate years where we will be forced to continue with both representations but we certainly think the subject is worthy of serious study.

Mr. Conlon's questions concerning the co-ordination of dual-element fuses with magnetic and manual motor controllers are good. A study of the time-current characteristics show that the dual-element fuse can furnish adequate short-circuit protection as well as back-up protection at the lower overloads so that if the controller fails, the motor will not be damaged. In such an application the controller will always open ahead of the dual-element fuse except when it needs protection.

Both the Hi-Cap and the Fusetron fuses are designed to carry 110-per-cent load when tested in the open. Both are low temperature, low resistance devices with extremely low watt loss so both will carry their rated load in an enclosure. Both would carry their ratings in an ambient of 125 F. Derating curves showing the effect of ambient on dual-element fuses are available. Ambients up to 200 F have little effect on the Hi-Cap fuse.

Every day, Hi-Cap fuses and Fusetron fuses are being co-ordinated with thermal breakers to protect equipment added to existing circuits, to protect additional circuits as well as to protect the circuit breakers themselves.

Mr. Kaufmann's comprehensive discussion is most welcome because it focuses attention on features which deserve further clarification. His statement that the allocation of only a portion of the fuse cartridge to the high-current element increases the internal stress under short-circuit interruption and requires a corresponding increase in strength to avoid explosion is a common fallacy many engineers assume to be true. Actually, as shown in Fig. 2, faster short-circuit operation with lower internal stresses is possible with the dual-element fuse because the operation at the lower overloads is isolated from the operation at short-circuit. Hence, the short-circuit links can be designed for optimum performance with minimum stresses under short-circuit conditions, and the overload device designed for optimum performance at the low overloads.

In like manner, Mr. Kaufmann's statement that the longer time delay imposes greater stress on devices which rely on the fuse for protection shows a lack of understanding of the dual-element fuse. Actually the longer time delay at the low overloads means that a smaller dual-element fuse can be used for the same load without experiencing useless blows on harmless overloads. Hence the fuse can be matched to the safe time-current characteristic of the equipment being protected, as explained in the paper.

As to the tests witnessed by the Electrical Testing Laboratories, in each test the fuse cleared the circuit without igniting readily inflammable material placed around the fuse with very little noise. In every test the fuse barrel remained intact and the ferrules remained in position on the fuse. Even though the performance of a circuit-breaker is considered satisfactory if the circuit-breaker can be rebuilt after the test, we would not consider the performance of a fuse as satisfactory if the fuse should explode in interrupting the current.

Mr. Kaufmann again is confusing circuit-breaker performance with fuse performance, in discussing the test circuit capacity. Actually the fuse is clearing the circuit in the first quarter cycle so that rms current is meaningless. The circuit breaker, on the other hand, permits the current to flow for several cycles so that rms values can be used. Even so, to avoid such discussion, we have given the Fusetron dual-element fuse a nominal interrupting rating of 100,000 amperes even though we are sure the fuse will clear higher currents.

Mr. Kaufmann does not feel that the co-ordination picture was developed completely because only the co-ordination under short-circuit conditions was discussed in detail. However, it is stated in the paper, "The time-current characteristics of the Fusetron dual-element fuses and the Hi-Cap fuses in ratings from 30 to 5,000 amperes, given in Fig. 6, show the co-ordination obtainable with these devices even though the time-lag of the Hi-Cap fuse is less than that of the dual-element fuse at the lower overloads."

To an engineer familiar with electrical protection Fig. 6 speaks for itself. However, for greater clarity, it might be well to state that, in general, proper co-ordination is obtained at the lower overloads if the curves do not cross each other, but a system study must be made for close co-ordination.

The curves of Fig. 6 show total time consisting of melting plus arcing time. As Mr. Kaufmann states, another set of curves showing "safe" time is required. The safe time is slightly less than the melting time. It is the time a given value of current can flow through a fuse without damaging the link. For proper co-ordination the total time of the smaller fuse must be less than the safe time of the larger one at all values of current. Curves showing the safe and total times of Fusetron and Hi-Cap fuses are available.

With reference to Mr. Kaufmann's comments on Fig. 7, he again is letting circuit-breaker custom affect his basic thinking. Actually the most common short circuit is a single-phase short circuit caused by two wires or conductors coming in contact. But power circuit breakers are 3-phase and a single-phase short circuit on a 3-phase breaker is too tough. Hence a 3-phase short circuit is selected for test. The Hi-Cap and Fusetron fuses are clearing single-phase short circuits with only one fuse in the circuit, which is the most severe condition. To obtain the value of assymmetrical rms amperes used in Fig. 7, the symmetrical amperes must be multiplied by 1.4.

Mr. Kaufmann condemns the use of the knife blade switch-and-fuse combination based on the report of work injuries in California by using totals instead of specific data. Table II of this report shows the breakdown of electrical work injuries by type of equipment and accident type in California during the fiscal year ending June 30, 1950. The table shows a total of 14 injuries in switches, externally operable, fused, ampere-rated; five injuries in switches, externally operable, fused, horsepower rated; and 16 injuries in circuit-breakers and motor starters. In 15 of the 19 switch injuries the workman was burned by the arc, three times he was hurt by a short circuit and once he was shocked as a result of his tool touching a live part. A breakdown of the circuit-breaker and motor starter injuries shows that seven times the workman was burned by the arc, three times hurt by a short circuit, four times he touched live parts, and twice he received a shock through equipment. From these figures it is seen that the statistics need more careful analysis before conclusions can be drawn concerning the relative safety of fused switches or circuit breakers.

I heartily agree with Mr. Kaufmann that proper co-ordination of the protective device with the equipment it is protecting is very important. We hope some day that the equipment suppliers will furnish safe time-current characteristics on their products, thereby simplifying the circuit designers' problem. In the mean time, the advent of the Fusetron and Hi-Cap fuse has done much to simplify the problem. It has not only established the maximum current which can flow under the most adverse conditions but, it also cuts the current off in a fraction of a cycle before it can destroy associated equipment and wiring.

Recent Developments in the Theory and Design of Electric Spark Machine Tools

E. M. WILLIAMS
MEMBER AIEE

J. B. WOODFORD, JR.
ASSOCIATE MEMBER AIEE

RICHARD E. SMITH
NONMEMBER AIEE

MACHINE tools in which material removal is effected by the action of successive electric sparks are commercially available in the United States, most western European countries, and are known to be in extensive use in the Soviet Union. As a substitute for industrial diamond tools in the machining of hard materials and as a means of performing machining operations which are otherwise impossible, electric spark machining has attracted extensive industrial interest. An earlier paper[1] described the basic process and a theory which explains the phenomena of spark machining. The purpose of this paper is to describe the results of further research into the fundamental processes and some design trends in the electric circuits of spark machine tools.

In the electric spark machining process a succession of high-current discharges are caused to pass between a work piece and a tool electrode having the cross section of the hole which it is desired to form in the work piece. Electrode and work piece are immersed in a dielectric fluid. The discharge is initiated by a breakdown or spark-over in this dielectric, which commonly takes place at a tool electrode to work piece separation of the order of thousandths of an inch and at voltages from 40 to 200 volts. There are two other types of electric machining process in commercial use, namely: 1. an oscillating electrode process in which discharges are contact-initiated, used in several models of tap-extractors; and 2. a grinding process in which a conducting electrolyte is used between a rotating flat electrode and a flat surface to be ground. In the absence of an accepted terminology in the prolific and confusing literature in this field, it is suggested that the three processes be distinguished as "spark-over" initiated discharge or "spark" machining, contact-initiated discharge machining and electrolytic grinding respectively. This paper deals only with the first of these three.

The basic phenomenon used in electric spark machining is the electrode erosion accompanying an electric discharge, particularly the intensified erosion at the anode resulting when discharge takes place in a liquid dielectric. This basic phenomenon has been studied for nearly two centuries, starting with the work of Priestly.[2] Bruma and Magat[3] have given an extensive bibliography of studies in this field. Numerous theories for the erosion mechanism have been advanced. The most recent theory is that outlined by Williams[1] in 1952, which attributes removal of anode material to electric field forces which arise from the extremely high current density at, about, and below the small area at the anode termination of the discharge. These forces act on the positive ions in the crystal lattice. The magnitude of the forces involved has been calculated[1] and found to be sufficiently great to account for the rupture of a portion of the anode surface. Experimental work cited gave results in accordance with predictions based on the electric field force hypothesis. In this paper a number of additional studies of the craters formed by single discharges will be described. These studies consisted of numerous experimental tests in which a single discharge, accurately controlled in duration and constant in current, was voltage-initiated between polished electrode specimens immersed in a hydrocarbon dielectric. After each test both anode and cathode electrodes were removed and the craters examined microscopically. Comparison of the variations of area, depth, and volume of craters with current magnitude and current duration and with chemical composition and physical properties of the electrodes has resulted in empirical relations governing the magnitude of anode and cathode material removal. It may be added, however, that these further experimental results are still in accord with the electric field force hypothesis.

When voltage is applied to an electrode gap of the type described, a discharge follows a delay which depends on the excess of applied voltage over that required for breakdown. A typical oscillographic trace of voltage is shown in Fig. 1. The time of discharge formation is of the order of 10 millimicroseconds (mms). Following formation the discharge voltage drop

Fig. 1. Oscillographic trace of voltage across gap between brass electrodes immersed in kerosene dielectric with electrode spacing of approximately 0.001 inch. The voltage is applied at the beginning of the trace and rises rapidly to 500 volts. After an initial delay of about 150 mms a discharge is formed and the voltage drops to a constant value of about 30 volts. Duration of the current discharge in this case is 1 microsecond and the current is 18 amperes

---

Paper 54-174, recommended by the AIEE General Industry Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing December 16, 1953.

E. M. WILLIAMS, J. B. WOODFORD, JR., and RICHARD E. SMITH are with Carnegie Institute of Technology, Pittsburgh, Pa.

Work described in this paper was supported by the Method X Company, affiliate of the Firth Sterling Corporation, Pittsburgh, Pa. This paper contains part of a dissertation by Richard E. Smith in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Carnegie Institute of Technology.

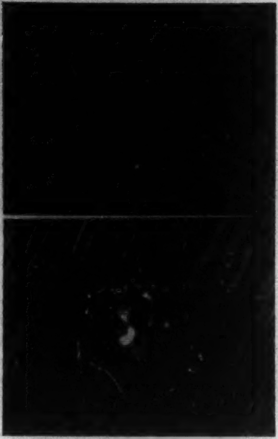


Fig. 2. Above, anode, and below, cathode, craters formed by a 1.2-microsecond discharge of 3 amperes in brass electrodes. Anode- and cathode-crater diameters are approximately 9.4×10^{-4} and 6.6×10^{-4} inch respectively

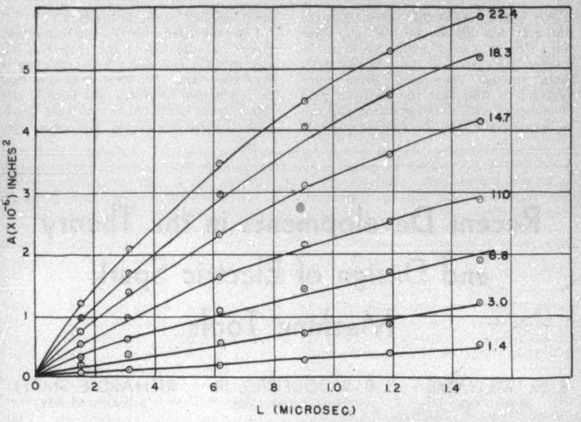


Fig. 3. Anode-crater area A as a function of current pulse duration in microseconds L for various values of discharge current in amperes. Areas are independent of anode- and cathode-electrode composition

is constant and of the order of 20 to 40 volts depending on electrode spacing. In the experimental tests to be described discharge durations were varied from 17 to 1,500 mms.

Fig. 2 shows typical anode and cathode craters produced by a single discharge. Cathode crater area is equal to or somewhat less than anode crater area. When the anode and cathode are of like material, the volume of the anode crater is greater than that of the cathode crater for pulse durations within the ranges employed in these experimental tests. Typical anode-to-cathode crater volume ratios for a pair of brass electrodes with a 15-ampere discharge are as follows:

Duration, Mms.	Ratio of Anode-Crater Volume to Cathode-Crater Volume
150	2.02
300	1.97
900	1.98
1,500	1.68

When anode and cathode materials are not alike, the anode-to-cathode crater volume ratios vary widely. It seems clear that crater size in an anode of a particular composition is independent of the cathode composition. On the other hand, the size of the crater in a cathode of a particular composition is radically affected by the composition of the anode. For instance, anode-crater tests with a 1,500-mms 11-ampere discharge for cathodes of brass and tungsten carbide gave anode-crater volumes as follows:

Cathode Composition	Anode-Crater Volume
Brass	11×10^{-9} inch^3
Tungsten carbide	11×10^{-9} inch^3

whereas corresponding measurements on the cathode show

Anode Composition	Cathode Volume, Cubic Inches
Brass	0.23×10^{-10}
Tungsten carbide	0.84×10^{-10}

Anode erosion is, therefore, a simpler phenomenon than cathode erosion. Furthermore, it can be demonstrated that anode erosion is closely related to current distribution in the surface of the anode body. For instance, anode craters were formed with 11-ampere discharges for two steel specimens as given in Table I.

Tensile strength (and conductivity) of these two specimens is similar, but craters formed in the ferromagnetic material are much smaller, particularly in the case of the shorter duration current. Since current skin effect is much more pronounced in the ferromagnetic specimen, again particularly for the shorter duration current pulses, while other physical phenomena are the same, it seems clear that the current density below the surface of the anode plays an important part in anode erosion. Other test results demonstrate the importance of tensile strength as compared with such factors as melting point, density, etc., in determining anode erosion, and are given in Table II.

Although anode-crater depths vary greatly from material to material, anode-crater area is a function only of current duration and magnitude. Figs. 3 and 4 show curves of area as a function of current duration and magnitude over a wide range. Crater area is particularly interesting because it yields discharge current density information.[4] An assumption that the current density is given by current divided by observed crater area may, however, yield fallacious information. Since crater area increases almost linearly with discharge duration, while depths remain approximately constant, it seems probable that current density is actually

Table I. Anode Craters for Two Steel Specimens

Steel Designation	Tensile Strength, Pounds per Square Inch	Current Duration, Mms.	Crater Volume, Cubic Inches
18-8 annealed (nonferromagnetic)	81×10^3 pounds per inch^2	150	8.4×10^{-11}
		1,500	2×10^{-10}
416 annealed (ferromagnetic)	91.4×10^3 pounds per inch^2	150	3.2×10^{-11}
		1,500	1.3×10^{-10}

Table II. Tensile Strength Compared with Other Factors

Anode Material	Melting Point, Degrees Fahrenheit	Tensile Strength, Pounds per Inch²	Density, Pounds per Inch³	Crater Volume, Cubic Inches	Current, Amperes	Duration, Mms.
Free-turning brass	1,625	61.2×10^3	0.307	9×10^{-10}	22	1,500
				3.6×10^{-12}	3	150
Grade-*B* hard phosphor bronze	1,922	59.6×10^3	0.321	9×10^{-10}	22	1,500
				3.6×10^{-12}	3	150
17ST aluminum	1,202	64.4×10^3	0.104	9×10^{-10}	22	1,500
				3.6×10^{-12}	3	150

Table III. Volumes of Anode Material

Material	Volume, Inch³ $\times10^{-9}$	Tensile Strength, Pounds per Square Inch (Ultimate)
2 *S* aluminum	0.32	18.7×10^3
Grade-*B* phosphor bronze	0.32	59.6×10^3
Free-cutting brass	0.32	61.2×10^3
17ST aluminum	0.32	64.4×10^3
Naval brass	0.27	75×10^3
8-8 steel	0.20	81×10^3
Tungsten carbide	0.11	150×10^3

constant and that increases in area with time are due to progressive displacement of a discharge path of constant area and density over the electrode surface as numerous small pieces are removed from the anode and cathode. In any case, the ratio of current to the entire crater area provides a means of calculating a minimum limit to the probable current density, the accuracy of which increases as the current duration is decreased. For very short pulses data have been obtained as follows:

Pulse Length, Mms.	Current	Apparent Current Density, Amperes per Inch²
17	18.3	110×10^6
17	22.4	106×10^6

Data at this short duration are obtained with some difficulty and the range of values of current is limited. However, a wider range of currents at greater duration, such as that shown in Fig. 5, readily demonstrates that apparent discharge current density is independent of the actual discharge current.

The total material removed from the anode electrode by a discharge of a given duration and magnitude provides an index of fundamental machinability by the electric spark process. Table III gives the volumes of anode material removed by a spark of 11 amperes and 1.5 microseconds duration for a few typical materials tested.

The experimental results of single discharge tests present a fairly simple picture of spark-cutting phenomena. Electric spark machine tools, however, operate through the action of successive discharges for the removal of more material than can be dislodged by a single discharge. In the majority of machines in use, the basic phenomena are complicated by the problems associated with deionization of the discharge path between successive discharges and the accumulation of machined "chips" in the discharge area.

Machine tools which operate by means of successive voltage-initiated discharges have taken several mechanical forms, usually analogous to the drill press or to the lathe. Machines for drilling and tapping are by far the most common. In the machines described by Lazarenko[5] and Teubner,[6] the successive discharges are derived from a charged capacitor in a relaxation circuit. A d-c source charges a capacitor through a charging resistance. The capacitor is in parallel with the machining gap. When the voltage between electrode (tool) and work piece rises to a sufficiently high value to puncture the dielectric, the capacitor discharges. Following the discharging the electrode gap deionizes and the capacitor resumes charging. Machines of this type have been termed "relaxation" machines by the authors. For small electrode to work piece spacings the repetition rate of discharges is high, and the individual discharge currents are small. As the electrode spacing increases, the repetition rate decreases but the individual discharge currents are increased. Fig. 6 shows an oscillograph trace of a typical charge-discharge train in a "Lazarenko" relaxation machine. Because the capacitor charging system is continuously connected, charging current may occasionally "follow over" the cutting discharge and result in the formation of a continuous arc which can be interrupted only by disconnecting the power source or by a radical withdrawal of the tool from the work piece. This difficulty is aggravated by the rapid initial rate of rise of the capacitor voltage which is characteristic of the resistance-capacitance circuit. Bruma and Magat[8] and others[7] have added to the charging circuit an inductor which has the effect of decreasing the initial rate of rise of gap voltage while increasing the rate near the end of the charging cycle. An oscillographic trace of a charge-discharge train in an inductor-

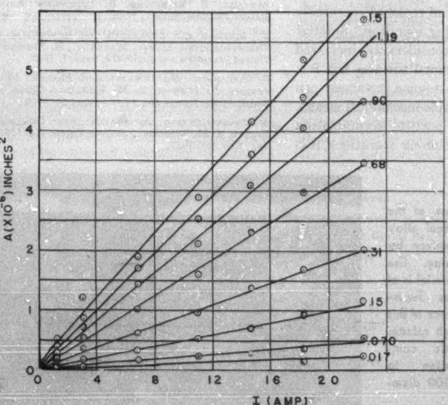


Fig. 4. Anode-crater area A as a function of current pulse magnitude, I amperes, for pulses of various durations in microseconds. Areas are independent of anode- and cathode - electrode composition

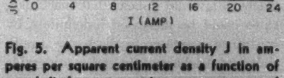


Fig. 5. Apparent current density J in amperes per square centimeter as a function of actual discharge current in amperes computed from data obtained with current pulses of 70 mms duration

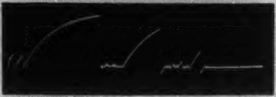


Fig. 6. Typical oscillographic trace of voltage across the discharge gap in a relaxation machine of the type described by Lazarenko. In this particular case a 62-microfarad capacitor is charged from a 110-volt d-c source through a 16.5-ohm resistor. The oscillographic trace length is approximately 10 milliseconds. Following a discharge immediately at the beginning of the trace the capacitor: 1. charged to about 30 volts; 2. discharged; 3. charged again to about 27 volts; 4. discharged; 5. charged to about 105 volts; 6. discharged; 7. was maintained at approximately zero because of a short circuit between electrodes; 8. charged to a voltage of about 92 volts, etc. The discharges are of such short duration as not to be clearly visible in long duration records such as this; however, the negative overshoot caused by the oscillatory character of the discharge is visible

resistor charging relaxation machine is shown in Fig. 7. Machines of this type have machining speeds of perhaps four times those of the basic Lazarenko type. As in the basic Lazarenko device, these machines are still characterized by a fluctuating discharge rate, associated varying voltage and discharge current, and inconsistency of machining tolerances. Bruma and Magat[3] have pointed out the advantages of a machine in which discharges are independently timed and controlled, but have disclosed no circuit details. A circuit which utilizes fixed amplitude currents from the secondary of a pulse transformer, the primary of which is periodically excited by pulses from a pulse-forming network, has been described by Williams.[8] Such machines which operate with independently-timed fixed-amplitude discharges provide a high degree of reproducibility and more reliable surface finishes. Since such circuits generally do not depend on discharge gap deionization for their timing, they may operate at higher discharge and cutting rates and are less erratic because arc follow-over is impossible. Such independently timed machines are invariably more costly in design and manufacture, however, than relaxation machines.

If the fundamental data on the effects of individual discharge are considered, it would appear that machining rates could be increased indefinitely by continued increases in discharge amplitude and discharge rate and considerable design effort has, in fact, been directed to this end. Independently timed electric discharge systems provide the means for achieving high discharge rates. The only apparent design problems involved are those of providing sufficient average power capacity for the electric discharge system and cooling capacity for the machining operation. Machines operating with discharge rates of up to 12,000 pulses per second, discharge currents up to 50,000 amperes, and average powers of 15 kw have been tested. In every case limiting values of cutting speed are reached before maximum discharge rates or average power limits are reached. This effect is known to be caused by the accumulation of electrode and work piece particles in the machining gap. This accumulation takes place in spite of forced circulation through the discharge gap of dielectric fluid under high pressures. In a typical tungsten-carbide die-machining operation about 0.1 cubic inch per minute of material is removed in a gap of about 0.001 inch with dielectric fluid circulated under a pressure of 30 pounds per inch^2 over an inside surface of 1 or 2 square inches. It seems inevitable that the dielectric in the machining gap should become heavily loaded with machined particles from both anode and cathode. This accumulation reduces machining speed at the anode drastically for reasons which are not completely understood, especially since the undesirable machining of the cathode electrode is not so drastically reduced. However, in deep cuts a characteristic, superficial rough "splatter" surface is superimposed on the machined workpiece. Fig. 8 shows the cross section of work piece machined surface and splatter-layer in a typical hard steel work piece. Machined particles apparently are driven by the explosive discharge forces, or other means, back against the machined surface to form a superficial layer. They may be removed and replaced repeatedly during the machining operation. Splatter layers result in considerably increased roughness in machined surfaces but fortunately, in at least some instances, can be removed by simple mechanical means.

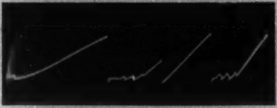


Fig. 7. Typical oscillographic trace of voltage across the discharge gap in a Lazarenko machine modified by the substitution of a resistor and reactor in the charging circuit. In this particular case a 62-microfarad capacitor is charged from a 110-volt d-c source through a circuit with an inductance of 0.2 henry and a d-c resistance of 1 ohm. The oscillographic trace length is approximately 20 milliseconds. Steps in the charge and discharge periods of this trace are similar to those of Fig. 6

Because of the particle accumulation effect in inside machining operations, it is believed that machining speeds now reached are very nearly the maximum attainable. In "outside" machining operations, analogous to grinding or milling, however, particle accumulation has not been a problem and possibilities exist for major increases over existing machining speeds. It is already possible, for instance, to exceed the maximum grinding speed obtainable by any other means in such hard materials as tungsten carbide.

References

1. Theory of Electric Spark Machining, **Everard M. Williams.** *AIEE Transactions*, vol. 71, pt. II, 1952, pp. 105–08.

2. An Account of Rings Consisting of all the Prismatic Colours, Made by Electrical Explosions on the Surface of Pieces of Metal, **Joseph Priestly.** *Philosophical Transactions*, Royal Society of London, London, England, vol. 58, 1768, pp. 68–74.

3. L'Usinage Par Electricite, **M. Bruma, M. Magat.** *International Mechanical Engineering Congress*, Milan, Italy, Oct. 1953.

4. Current Densities in the Cathode Spots of Transient Arcs, **S. M. Somerville, W. R. Blevin.** *Physical Review*, New York, N. Y., vol. 76, no. 9, 1949, p. 982.

5. Procede et Machine pour le Travail de Metaux, **N. Lazarenko, B. Lazarenko.** Swiss Patent *257,468*, April 1, 1949.

6. Method and Apparatus for Electrically Disintegrating Metal Material, **E. Teubner.** United States Patent *2,650,979*, Sept. 1, 1953.

7. Apparatus for Electrically Eroding Materials, **E. Teubner, E. M. Williams.** Union of South Africa Patent No. *14,062*, May 26, 1952.

8. Spark-Cutting Apparatus and Control System, **E. M. Williams.** Belgian Patent No. *515,461*, Nov. 29, 1952.

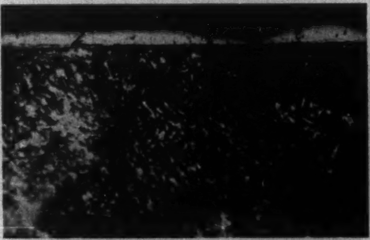


Fig. 8. Cross section of the surface of a hard steel alloy finished to a flat surface by electric spark machining. The main body of the material is of austinitic structure. Superposed on the flat surface of this main body is a rough splatter layer of martinsitic composition. Magnification in this photograph is 500 diameters

Discussion

M. Magat (Laboratoire de Chimie Physique de la Faculté des Sciences de Paris, France): On all essential points we entirely agree with the results of the authors.

There are however a few details where we differ, particularly concerning the fundamental machinability of different materials. Although it is not impossible that the discrepancy arises from the difference in the criterion used, I think it worthwhile to state briefly our results.

We took as the "machinability index" not the volume of the elementary crater but the number of watt-hours necessary to remove 1 gram of the material under standard conditions and the amount of material in grams removed per unit time P_{rel}, taking brass as standard. As can be seen from Table IV, the results obtained by the two methods are closely parallel and agree rather well with the results of Lazarenko[1] (column I) and of Benedicks[2] (column II). So striking is the discrepancy with the results given by the authors that we recalculated in grams, again taking brass as unity (column III).

As can be seen, on our scale, steel is slightly more easily and aluminium much more easily machinable than tungsten carbide, in contradiction with results of the authors.

Much the same, if we agree that the crater volume of the cathode, which is a measure of the "wear" of the tool electrode, very strongly depends on the nature of the anode, we found, contrary to the authors, that the rate of cutting of a given anode also strongly depends on the nature of the cathode. For instance, we found that rate of material removal from tungsten carbide (Carboram *FG*) is four times as fast if the cathode is copper than if it is in aluminium. It is of course possible that brass and tungsten carbide give identical results, but we think it is rather a coincidence than a general rule. There again we must emphasize that the disagreement may be due to the tests used: crater volume versus weight removed per unit time.

References

1. See reference 5 of the paper.
2. **Benedicks.** *Arkiv för Mathematik Astronomi-och Fysik*, Stockholm, Sweden, vol. 8, no. 7, 1912.

M. S. Bruma (Laboratoire de Chimie Physique de la Faculté des Sciences de Paris, France): There is a serious fundamental limitation, for a given electrode size, to the inference that "machining rates could be increased indefinitely by increases in discharge amplitudes and discharge rates," since rapid increases in the skin-effect losses would inevitably be caused by the variations involved in these factors.

The skin-effect losses, P watts, in a cylindrical electrode of length l centimeters and radius a centimeters, through which flows rectangular current pulses of peak value I_p amperes, duration L seconds, and repetition frequency f_r, is given by

$$P=2.7\times10^{-8}[L^{1/2}f_rI_p{}^2l/a] \tag{1}$$

derived from $P=R'I_{\text{rms}}{}^2$, with R', the skin effect resistance, calculated for a frequency $(2L)^{-1}$ and $I_{\text{rms}}=(Lf_r{}^{1/2}I_p)$ for rectangular pulses.

For a 10-centimeter-length copper electrode of radius 1 centimeter, with $L=0.5\times10^{-6}$ second, $f_r=1{,}200$, and $I_p=5{,}000$ amperes, the skin-effect losses calculated from equation 1 would be 5.7 watts. If it is desired to increase the discharge amplitude I_p and the discharge rate f_r, each by a factor of 10, then, for the same electrode, the losses would be multiplied by a factor of $(10)^2\times10=1{,}000$, rising to the uncomfortable value of 5.7 kw. This is only part of the losses occurring in the discharge circuit; total losses should include the workpiece and the leads.

However, equation 1 shows that by increasing simultaneously the electrode radius, a compromise may be found between high machining rates and reasonable losses; e.g., with a radius of 10 centimeters, the aforementioned electrode losses would be reduced to the more realistic value of 0.57 kw.

Another approach to the problem is to consider the sparks between the electrode and workpiece as a perfect transformer of electrical to mechanical energy, and define an "equivalent primary resistance" R, so that $RI_{\text{rms}}{}^2$ represents a quantity proportional to the rate of metal removed. Furthermore, considering the spark discharge as an avalanche of electrons, it appears that R must be of a nonlinear nature, proportional to l/I_{rms}. This has been confirmed in our experiments, which gave fairly constant values for the ratio $(V/I_{\text{rms}}-V)$ to (machining rate) in the range from 10 to 100 amperes, for given electrode, workpiece, and dielectric fluid. The total electric power supplied to the discharge circuit may then be divided between two resistors in series

$$P_t=R_LI_{\text{rms}}{}^2+(k/I_{\text{rms}})I_{\text{rms}}{}^2$$

R_L being responsible for the total losses and $k/I_{\text{rms}}=$ the "machining" component. Now, by continuous increase in I_{rms} it is clear that the limit value of the ratio $\Delta V/\Delta P_t$ will be zero, since V is proportional to I_{rms} and the losses are proportional to I_{rms}.

It becomes apparent that, for each electrode size, there is an optimum power carrying capacity determined by skin effect considerations which has a "saturating" effect upon the machining rates. The curve: $V-P_t$ will be similar to the well known $B-H$ curve for a magnetic material, the knee of the curve giving the optimum value for V (machining rate) and P (power input).

Table IV. Machinability Indexes

Metal or Alloy	No. of Watt-Hours to Remove 1 Gram	Amount of Material, Grams Removed per Unit Time, P_{rel}			
		I	II	III	IV
Lead	3	4.6	6.8	24.6	
Nickel	11	1.3	0.71		
Aluminum	16	1	1.5	(1)	0.3
Brass	16	(1)	(1)		(1)
Titanium	43	0.55			
Molybdenum	46	0.47	0.42	0.49	
Steel		0.5			0.6
Tungsten carbide		0.35			0.6

Carroll R. Alden (Ex-Cell-O Corporation,* Detroit 32, Mich.): A study of the references appended to the paper reveals a situation which is perhaps unique in the annals of technical development. In this instance a usable electric circuit and certain desirable results of its use which, incidentally are obtainable by no other known means, were known for nearly 200 years before the first improvement was made in the apparatus.

It seems in order, therefore, to congratulate a relatively unknown worker, Edmund E. Teubner, for conceiving and applying to the basic spark cutting circuit, first presented by Priestly in 1768, an improvement (which was published by Bruma and Magat in 1953).

Likewise, it seems in order to congratulate the Firth Sterling Steel and Carbide Company of Pittsburgh for recognizing the commercial possibilities of this improved apparatus and for creating their wholly owned subsidiary, The Method X Company, to continue the development thereof in the laboratories of the Carnegie Institute of Technology and elsewhere.

It perhaps is not too late to congratulate Dr. Williams for his searching analysis which enabled him to present, before the Institute in 1952, the nature of the forces by which particle dislodgment from the anode is effected.

Finally, it is our desire to congratulate the authors of the present paper for advancing this as yet little known art by the present contribution of data and photomicrographic records.

It is understandable that, at this stage of the development, engineers are largely devoid of the experiences and interests promotive of a more animated discussion of the subject matter. This situation undoubtedly will alter as increasing interest justifies the allocation of additional time to the presentation of papers on this subject, and as the presently existing and rapidly growing body of knowledge pertaining to this very important subject can be more comprehensively presented.

The long-known arc-cutting procedure by which an electrified tool may effect the cutting of an oppositely electrified workpiece in the absence of physical contact between the electrodes has always been limited in its utility by the deleterious effect of ambient temperatures and the absence of that degree of precision which is necessary for many important uses. The apparatus and process to which the present paper pertains is not subject to such limitations.

The rapidly growing importance of the ability precisely to machine the heat resistant, and heretofore substantially unmachinable, materials (required for the blades of improved internal-combustion turbines) would assure rapid and comprehensive development of this apparatus, even if all other incentives were wholly absent.

---

* The Ex-Cell-O Corporation is the exclusive licensee for the United States under inventions owned by The Method X Company, who sponsored this research.

E. M. Williams, J. B. Woodford, Jr., and **Richard E. Smith:** The discussion by Dr. Magat is of considerable interest since it reports the results of studies of machining operations in which a very large number of successive discharges are applied to an operation in an electric-spark machine tool as opposed to the single discharge data given by the authors. The rate of cutting in such an operation depends on a number of factors, in addition to the amount of anode material removed by a particular discharge. It is the authors' opinion that the differences reported by Dr. Magat are due to particle accumulation and electrode-clearance phenomena. For instance, it has been found by the authors that the electrode to workpiece separation when aluminum is used as an electrode is very much less for a given critical spark-over voltage than for the other electrode materials discussed. When aluminum is used as an electrode in a machine tool, the smaller clearances result in much more rapid clogging of the machining gap with both electrode and workpiece material, and consequently a slower over-all process of machining. The difference in spark-over voltage is believed to be due to a difference in electron-emission phenomena as a result of the oxide coating on the aluminum electrode.

In studying the difference between copper and brass electrode materials, the authors have found that wide variations in electrode wear occur as a function of discharge-current duration. For very short duration discharges, wear of copper electrodes has been very much less than that of brass. For longer discharges, little difference in wear is encountered. This effect of discharge duration has been attributed by the authors to skin effect, which is more pronounced in copper than in brass, and particularly so with the shorter discharges. The accumulation of particles is a major effect in determining machining speed; electrode particles are often produced in greater volume than workpiece particles, so that the cathode material used affects over-all machining rate, despite the fact that cathode material does not affect the volume of workpiece removed by a single discharge.

Dr. Bruma is correct in pointing out the limitations imposed by energy losses in the electrode in an attempt to increase machining speed by increases without limit in the current of each single discharge. This limitation does not apply, however, to increases in machining rate effected by increases in discharge repetition rate. Increase in repetition rate is a more attractive approach to the problem of increased speed, since it results in finer surface finish and greater accuracy.

There is, of course, an ultimate limitation on machining rate imposed not by efficiency of a single discharge, but by the steady dissipation of heat. In a machining operation with electrode and workpiece of a given size, the power which can be employed is limited by the cooling capacity and, as in any engineering operation, increased power can be used when cooling is improved. The authors' conclusions concerning great increases in machining rates can, therefore, only be applied to machining operations carried out over larger and larger surfaces, and Dr. Bruma's point is well taken.

We are grateful to C. R. Alden for his expression of the point of view of a machine-tool user of the electric-spark process.

The Application of Short-Time Memory Devices to Compensator Design

D. J. FORD
ASSOCIATE MEMBER AIEE

J. F. CALVERT
FELLOW AIEE

IN FIG. 1 is shown a compensator cascaded with a control system. The combination will be called the compensated system. In this paper the transfer function of the control system is assumed to be linear and that for the compensator is at least piecewise linear. Each compensator discussed in the paper has incorporated, within its structure, a delay line or short-time memory unit, Fig. 2, into which information is inserted and from which the same information is extracted at various short intervals of time thereafter. The extracted information at each location is multiplied by a suitable constant and the results are summed to form $r_x(t)$.

Referring again to Fig. 1, the purpose of a compensator is to compel the response $q(t)$ to approximate closely some function of the command or input signal $v(t)$. In this paper the command will be sinusoidal in form and the purpose of the compensator will be to make the transfer function of the compensated system $q(t)/v(t)$ have a constant amplitude and linear phase shift over a range of values of angular frequency $0 \leqq \omega \leqq \omega_1/2$. In general, the transfer function for the control system alone $q(t)/r(t)$ may approximate these desired characteristics for limited ranges of angular frequencies. For the linear control systems studied, it will be shown that with appropriately designed compensators this frequency range, in theory at least, can be greatly, if not indefinitely, expanded.

General Design Objectives

From Fig. 1, the basic requirement upon the compensator transfer function is that over the frequency range of interest

$$\frac{D_x(s)}{N_x(s)} \approx \frac{D(s)}{N(s)} \tag{1}$$

List of Symbols

TIME-ORIENTED QUANTITIES

f = frequency of command function
s = complex variable of Laplace transformation calculus
t = instantaneous time
T = a period of time in units of t
$T_d = 2T_K$ = total time delay possible within the short-time memory unit
T_k = the kth time delay as indicated in Fig. 2
$\frac{dB_a(T)}{dT}, \frac{dB_b(T)}{dT}$ = the multipliers per unit of time T along the delay line, as utilized in the modified Fourier integral compensator (FIC) design
$\omega = 2\pi f$ = angular frequency
ω_1 = twice the angular frequency corresponding to the upper limit of compensation for the modified Fourier series compensator (FSC) unit

VARIABLES WHICH ARE FUNCTIONS OF TIME

$q(t)$ = output, or response, either the controlled or indirectly controlled variable
$r(t)$ = output from a compensator element and the reference input to the next element to the right in Fig. 1
$r_x(t)$ = output from a compensator element and the reference input to the compensator element to the right in Fig. 1
$v(t) = |v|\epsilon^{j\omega t}$ = command function, a complex function of time

TRANSFER FUNCTIONS

$\frac{N(s)}{D(s)}$ = complex transfer function of the control system, where $N(s)$, $D(s)$ are assumed to be polynomial functions of s

---

Paper 54-124, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing November 25, 1953.

D. J. FORD and J. F. CALVERT are with Northwestern University, Evanston, Ill.

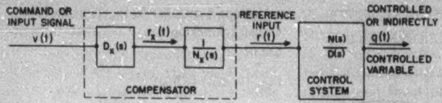


Fig. 1 (above). General diagram of compensated system

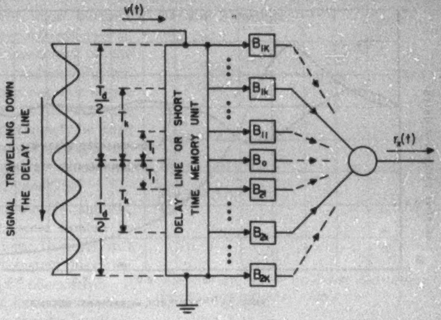

v(t)
SIGNAL TRAVELLING DOWN THE DELAY LINE
T_d/2
T_k
T_1
DELAY LINE OR SHORT TIME MEMORY UNIT
B_1K
B_1k
B_11
B_0
B_21
B_2k
B_2K
r_x(t)

Fig. 2 (right). Principles for forming the D(s) network of the compensator

B_{1k}, B_{2k}, etc. = REAL CONSTANTS
T_1, T_k, etc. = DISCRETE TIME DELAYS

$\frac{D_x(s)}{N_x(s)}$ = complex transfer function of the compensator

$\frac{N(\omega)}{D(\omega)}$ = steady-state sinusoidal transfer function of the control system

$\frac{D_x(\omega)}{N_x(\omega)}$ = steady-state sinusoidal transfer function of the compensator

$A_1(\omega) = \epsilon^{+j\omega T_d/2} A(\omega)$

$jB_1(\omega) = \epsilon^{+j\omega T_d/2} jB(\omega)$

$a(\omega)$, $jb(\omega)$, $A(\omega)$, $jB(\omega)$ = functions of ω, defined where introduced

$Y_a(\omega)$, $jY_b(\omega)$ = steady-state sinusoidal transfer functions of a portion of the general compensator structure

PARAMETERS

$b_0, \ldots, b_k, \ldots, b_K$; $B_{a0}, \ldots, B_{ak}, \ldots, B_{aK}$; $B_{b1}, \ldots, B_{bk}, \ldots, B_{bK}$; $B_{11}, \ldots, B_{1k}, \ldots, B_{1K}$; $B_{21}, \ldots, B_{2k}, \ldots, B_{2K}$ = real numbers and are defined either directly or by usage when introduced

SUBSCRIPTS

$0, 1, \ldots, k, \ldots, K$ = continuous sequence of integers
d = subscript to indicate a relation to the length of the short-time memory unit
g = positive number (an integer)
l = a real number permitted to become infinite in the limit (FIC unit)
m = positive odd number (an integer)
n = positive even number (an integer)
N = subscript to indicate relationship to $N(\omega)$

Consider the control system terms, $N(s)$, $D(s)$ to be polynomial functions of s. Let it be assumed that the compensator is to be designed in two parts, as indicated by the two blocks within the compensator of Fig. 1. Conventional network theory can be utilized to produce a transfer function in terms of passive circuit elements of the form $1/N_x(s)$, i.e., a transfer function having complex frequency poles and no zeros. By this procedure $N_x(s)$ might be made equal to $N(s)$. Because this design procedure appears to exist already, the remaining discussion will be concerned exclusively with the design of the $D_x(s)$ device. Thus, it is required that $D_x(s)$ approximate the polynomial $D(s)$. The problem becomes that of designing a network, the transfer function of which contains zeros and no poles. It is shown in the Appendix that the general network of Fig. 2 has this type of transfer function.

In Fig. 2 the short-time memory device itself may be an electrical delay line, or a magnetic or electrostatic recording device. In the delay line, the command signal moves through the tapped line and is dissipated in the terminating impedance. With the magnetic recorder, the command signal would be introduced into the tape, and the tape arranged as a moving continuous belt or disc. In this manner, the command signal would be available for recovery until such time as erasure becomes necessary to provide storage space for new signal information. To simplify the subsequent discussion, the memory device will be referred to as a delay line, although the type of delay or memory unit employed does not alter the basic theory.

Historical Background

Before entering upon the theoretical discussion for this paper, three historically related studies will be reviewed briefly. In 1940 H. E. Kallman[1] described a transversal filter network which was of the same general form as the delay-type network of Fig. 2. However, the purpose of that particular work was to produce a filter which had the required amplitude characteristic and which, within itself, produced a linear phase lag response. Apparently no attempt was made to establish the necessary design procedure or to utilize that network as a compensator for operation in combination with a control system so as to produce prescribed resultant over-all amplitude and phase characteristics. The work of A. M. Hopkin[2] in 1951 described the use of an amplified command signal plus two delayed and amplified signals to control a nonlinear second-order system under the conditions which would be imposed by a step function command signal. The design of a delay-type compensator referred to as the SCC device was described in 1952.[3] It was intended for use with linear control systems of any finite order when the command signal could be approximated by polynomial functions of time.

This leads to the present paper, which provides a steady-state frequency analysis that may be employed in designing compensators to go with control systems of any finite order and which, through this approach, presents three new design types.

Delay-Type Compensators Designed on a Frequency Response Basis

For steady-state sinusoidal conditions, the Laplace Calculus complex variable s is replaced in the transfer function expressions by $j\omega$, where ω is the angular frequency. The compensator transfer function $D_x(\omega)$ then possesses real and imaginary parts which will be designated as $A(\omega)$ and $jB(\omega)$ respectively, while the control system transfer function $D(\omega)$ will be called $a(\omega)+jb(\omega)$. The basic design equation for $D_x(\omega)$ over a given frequency range will be

$$D_x(\omega) = A(\omega) + jB(\omega) \approx a(\omega) + jb(\omega) = D(\omega) \quad (2)$$

It will be found that the terms in the polynomial $a(\omega)$ contain only even powers of ω, while the terms in the polynomial $jb(\omega)$ contain only odd powers of ω. In the Appendix it is shown that $A(\omega)$ contains a linear phase delay term multiplied by cosine terms in ω, while $jB(\omega)$ contains the same linear phase delay term multiplied by sine terms in ω. See equations 23 and 25. Then, over a given range, $0 \leq \omega \leq \omega_1/2$, the polynomial $a(\omega)$ may be represented by a Fourier cosine series and $jb(\omega)$ by a Fourier sine series. The coefficients of the cosine terms in $A(\omega)$ may be chosen to match, term for term, with a finite number of those coefficients in the representation of $a(\omega)$. A similar matching procedure may be employed to make the coefficients of $jB(\omega)$

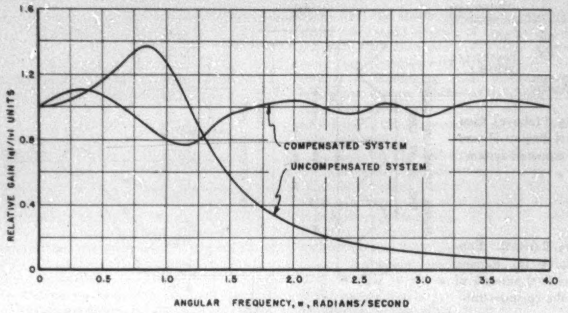


Fig. 3. Amplitude response for an FSC compensated second-order system. K=5, T$_d$=14 seconds. (Uncompensated system parameters: damping ratio=0.4, undamped natural angular frequency=1.0 radian per second)

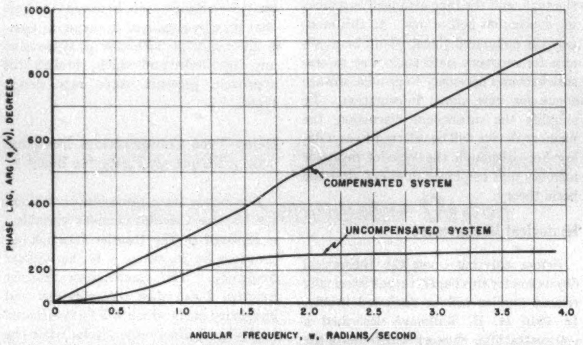


Fig. 4. Phase response for an FSC compensated second-order system. K=5, T$_d$=14 seconds. (Uncompensated system parameters: damping ratio=0.4, undamped natural angular frequency=1.0 radian per second)

match with a finite number of those in the Fourier representation of $jb(\omega)$. The foregoing is basic to the first compensator design which is to be described. Then, by appropriate alterations in the theory, the possibilities of the two subsequent designs will be explained.

The first design, FSC, will be described with reference to Fig. 2. The sinusoidal command signal $v(t)$ enters the delay line and travels past the various tap points. The energy of this signal is dissipated in the terminating impedance. However, it is observed that the signals appearing at the various tap points have been delayed in time with respect to the incoming command signal. Thus, the signal at the center tap has been delayed by $T_d/2$ seconds and, consequently it possesses a phase lag of $\omega T_d/2$ radians with respect to the command signal.

If each of the signals from two tap points symmetrically disposed about the center of the delay line were to be multiplied by $B_a/2$ and then added, the transfer function of the combination would become

$$Y_a(\omega)=\epsilon^{-j\omega T_d/2}(B_a/2)(\epsilon^{-j\omega T_k}+\epsilon^{-j\omega T_k}) \quad (3)$$

$$=\epsilon^{-j\omega T_d/2}B_a\cos\omega T_k \quad (4)$$

If, instead, the signals from these same two tap points were to be multiplied by $B_b/2$ and then the signal from the lower tap subtracted from that of the upper, the following transfer function would result

$$jY_b(\omega)=\epsilon^{-j\omega T_d/2}jB_b\sin\omega T_k \quad (5)$$

Proceeding in approximately this fashion, it is possible to develop a single network such that the resultant transfer function becomes the sum of terms of the form of equations 4 and 5. In fact, the resultant transfer function for $D_x(\omega)$ may be written as

$$D_x(\omega)=A(\omega)+jB(\omega)$$

$$=\epsilon^{-j\omega T_d/2}\left[B_{a0}/2+\sum_{k=1}^{k=K}B_{ak}\cos\omega T_k+ j\sum_{k=1}^{k=K}B_{bk}\sin\omega T_k\right] \quad (6)$$

Here, it should be mentioned that a more complete mathematical development is given in the Appendix. With reference to the last result, also see equations 26 through 29.

As stated earlier, the components $a(\omega)$ and $jb(\omega)$ of the control system transfer function may be expressed as a Fourier cosine and Fourier sine series respectively. The coefficients $B_{a1}\ldots B_{aK}$ and $B_{b1}\ldots B_{bK}$ of equation 6 are those to be matched with the coefficients of the corresponding terms in the combined Fourier representation of $a(\omega)+jb(\omega)$.

Figs. 3 and 4 give the calculated amplitude and phase characteristics of a second-order system before and after compensation by an FSC unit designed in the foregoing manner.

It was thought that better performance might be obtained if the number of taps along the delay line were allowed to increase without limit (within the fixed finite total length, T_d). This theoretical possibility was investigated, and the compensator so formed is called the modified Fourier integral compensator (FIC) because of the Fourier integral-like expressions which appear in the mathematical development given in the Appendix. The FIC network, when designed for use with a second-order control system, produced resultant amplitude and phase characteristics similar to those appearing in Figs. 3 and 4. Thus, no very significant differences appeared in the performance of the compensated systems when one incorporated an FSC and the other an FIC unit, and the same control system was employed in each.

Both FSC and FIC units gave the compensated system an oscillatory type of amplitude response. This is the result of trying to fit relatively low-degree polynomials with trigonometric functions. An additional shortcoming was that it appeared necessary to accept a rather large linear phase lag. Although this lag might be tolerated in some instances, many control systems require a smooth amplitude response and a relatively small linear phase lag. The final compensator to be considered met both of these specifications.

The last device to be discussed is called the modified Taylor's series compensator (TSC). If the polynomial functions $a(\omega)$ and $b(\omega)$ are expanded in Taylor's series about the origin, they yield themselves plus an infinite number of terms with zero coefficients. If in equation 6 the $\cos \omega T_k$ and $\sin \omega T_k$ terms are all expanded in Taylor's series with respect to the origin, a finite number of infinite series will result. Then, the sum of the coefficients of a given power of ω in the infinite series representing $D_x(\omega)$ may be equated to the coefficient of the corresponding power of ω in Taylor's series representation of $D(\omega)$. If, in equation 6, K is chosen sufficiently large, all of the nonzero and a finite number of the zero coefficients in the Taylor's series representation of $D(\omega)$ can be matched exactly by the Taylor's series representation of $D_x(\omega)$. Again, the mathematical details of this procedure are given in the Appendix.

The TSC produced greatly improved results, as evidenced by the amplitude and phase characteristics for a second-order control system with which is cascaded a TSC unit. See Figs. 5 and 6. It appears that constant unity amplitude response with an arbitrarily small linear phase lag might be obtained for any desired finite frequency range.

Observations and Conclusions

The historical review indicates that different investigators have established different criteria as the basis for the design of delay-type compensators. The present paper assumes that the steady-state amplitude and phase response characteristics of the compensated system are specified. Even though designed upon different bases, it may still be useful to compare all of the compensators mentioned on the basis of their ability to produce unity amplitude and linear phase response over a broad frequency range. On this basis, the various systems referred to may be listed in the order of increasingly desirable performance: the Kallman transversal filters, FIC, FSC, SCC[3], and TSC.

The performances of the FSC and FIC compensated systems were limited by theoretical considerations. However, with the TSC compensated system theoretical limitations were almost nonexistent. It seems probable that if, for control systems which have low-degree transfer functions, TSC units are employed, then either saturation in the control system or design complexities in the structure of the compensator will provide the primary limitations on the results which can be achieved.

In all cases it will be noted that the quality of the performance depends upon the ability to fit desired polynomials in ω with $D_x(\omega)$. Thus, the performance is vitally dependent upon the arrangement of the taps and multipliers used with the delay line. It is conceivable that there may exist high-degree polynomials for which the FSC or the FIC method might yield better theoretical results than the TSC.

Appendix

Throughout the mathematical treatments the time variables are complex, as described in the list of symbols.

General Relations

In Fig. 1

$$q(t) = \frac{N(\omega)}{D(\omega)} r(t) \tag{7}$$

where

$$N(\omega) = a_N(\omega) + jb_N(\omega) \tag{8}$$

$$D(\omega) = a(\omega) + jb(\omega) \tag{9}$$

$$a(\omega) = b_0 - b_2\omega^2 + b_4\omega^4 - \ldots + (-1)^{n/2} b_n\omega^n \tag{10}$$

$$jb(\omega) = j[b_1\omega - b_3\omega^3 + b_5\omega^5 + \ldots + (-1)^{(m-1)/2} b_m\omega^m] \tag{11}$$

where n and m are even and odd positive integers respectively. Then

$$q(t) = \frac{D_x(\omega)}{N_x(\omega)} \frac{N(\omega)}{D(\omega)} v(t) \tag{12}$$

It will be assumed practicable to construct a passive network for the representation of

$$\frac{1}{N_x(\omega)} \approx \frac{1}{N(\omega)} \tag{13}$$

For a follower system it is desirable that

$$D_x(\omega) \approx D(\omega) \tag{14}$$

The physical realization of $D_x(\omega)$ is much more difficult, and the discussion of various methods for achieving this are presented in the following.

FSC

The terms $a(\omega)$ and $jb(\omega)$ are polynomials (see equations 10 and 11). In this section each of these polynomials will be approximated by a Fourier series over an angular frequency range, $0 \leq \omega \leq \omega_1/2$ and then methods will be developed for an approximate physical representation. Let

$$a(\omega) = B_{a0}/2 + \sum_{k=1}^{k=\infty} B_{ak} \cos (2\pi k/\omega_1)\omega \tag{15}$$

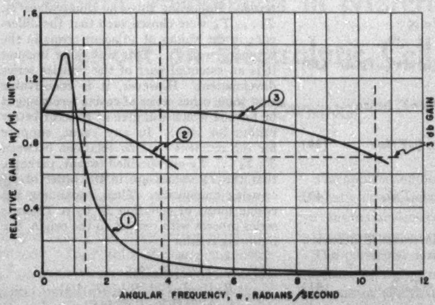


Fig. 5. Amplitude response for a TSC compensated second-order system. 1 = uncompensated system; 2 = compensated system, $T_d = 1$ second, $K = 1$; 3 = compensated system, $T_d = 1$ second, $K = 2$. (Uncompensated system parameters: damping ratio = 0.4, undamped natural angular frequency = 1.0 radian per second)

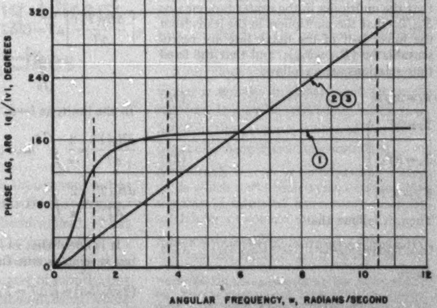


Fig. 6. Phase response for a TSC compensated second-order system. 1 = uncompensated system; 2 = compensated system, $T_d = 1$ second, $K = 1$; 3 = compensated system, $T_d = 1$ second, $K = 2$. (Uncompensated system parameters: damping ratio = 0.4, undamped natural angular frequency = 1.0 radian per second)

$$jb(\omega)=j\sum_{k=1}^{k=\infty} B_{bk}\sin(2\pi k/\omega_1)\omega \tag{16}$$

$$D(\omega)=a(\omega)+jb(\omega) \tag{17}$$

where

$$B_{a0}=(4/\omega_1)\int_0^{\omega_1/2}a(\omega)d\omega \tag{18}$$

$$B_{ak}=(4/\omega_1)\int_0^{\omega_1/2}a(\omega)\cos(2\pi k/\omega_1)\omega d\omega \tag{19}$$

$$B_{bk}=(4/\omega_1)\int_0^{\omega_1/2}b(\omega)\sin(2\pi k/\omega_1)\omega d\omega \tag{20}$$

Refer now to Fig. 2 and consider, for the moment, only the solid-line connections at the output from the short-time memory unit. Assume that the multipliers B_{1k} and B_{2k} are each equal to $(B_{ak}/2)$, and that the outputs are summed. Under these conditions

$$r_x(t)=\epsilon^{-j\omega T_d/2}\left(\frac{B_{ak}}{2}\right)(\epsilon^{+j\omega T_k}+\epsilon^{-j\omega T_k})v(t)$$
$$=\epsilon^{-j\omega T_d/2}B_{ak}\cos\omega T_k v(t) \tag{21}$$

If a similar procedure were followed for all output pairs symmetrical about the center tap, and if the multiplier on the center tap is $B_{a0}/2$, then

$$r_x(t)=A(\omega)v(t) \tag{22}$$

where

$$A(\omega)=\epsilon^{-j\omega T_d/2}\left[\frac{B_{a0}}{2}+\sum_{k=1}^{k=K}B_{ak}\cos\omega T_k\right] \tag{23}$$

Next, if the multipliers B_{ak} were replaced by multipliers B_{bk} and if the total output were the sum of those from the upper half of the delay line in Fig. 2 less the sum of those from the lower half, then

$$r_x(t)=jB(\omega)v(t) \tag{24}$$

where

$$B(\omega)=j\epsilon^{-j\omega T_d/2}\sum_{k=1}^{k=K}B_{bk}\sin\omega T_k \tag{25}$$

Next, for Fig. 2 assume that the multipliers in the leads from the upper half of the delay line are representable by $(B_{ak}+B_{bk})$; that the multiplier in the center tap remains $B_{a0}/2$; that the multipliers in the leads from the lower half of the delay line are representable by $(B_{ak}-B_{bk})$; and that the fixed time relations are as follows

$$T_1=2\pi/\omega_1 \tag{26}$$

$$\left.\begin{array}{l} T_2=2T_1 \\ \vdots \\ T_k=kT_1 \\ \vdots \\ T_K=KT_1 \end{array}\right\} \tag{27}$$

Then, it follows that

$$r_x(t)=[A(\omega)+jB(\omega)]v(t)=D_x(\omega)v(t) \tag{28}$$

where

$$D_x(\omega)=\epsilon^{-j\omega T_d/2}\left[B_{a0}/2+\sum_{k=1}^{k=K}B_{ak}\times\right.$$
$$\left.\cos(2\pi k/\omega_1)\omega+j\sum_{k=1}^{k=K}B_{bk}\sin(2\pi k/\omega_1)\omega\right] \tag{29}$$

A comparison of equation 29 with equations 15 through 17 shows that $D_x(\omega)$ approximates $D(\omega)$, but that part of that approximation involves a linear phase lag described by the term $\epsilon^{-j\omega T_d/2}$

FIC

Suppose that it is desired to keep the total time lag of the delay line at the value T_d, but that it is desired to insert many more taps within the length of this unit to widen the frequency range to $0\leq\omega\leq l\omega_1/2$. Here l is a real number which in the following discussion will be allowed ultimately to increase without limit. For convenience in this work, equation 29 will be rewritten as follows

$$D_x(\omega)=\epsilon^{-j\omega T_d/2}[A_1(\omega)+jB_1(\omega)] \tag{30}$$

where

$$A_1(\omega)=B_{a0}/2+\sum_{k=1}^{k=K}B_{ak}\cos(2\pi k/'\omega_1)\omega \tag{31}$$

$$jB_1(\omega)=j\sum_{k=1}^{k=K}B_{bk}\sin(2\pi k/l\omega_1)\omega \tag{32}$$

For all values of $k=0, 1 \ldots K$

$$B_{ak}=(4/l\omega_1)\int_0^{l\omega_1/2}a(\omega)\cos(2\pi k/l\omega_1)\omega d\omega \tag{33}$$

And for values of $k=1, 2 \ldots K$, it is written that

$$B_{bk}=(4/l\omega_1)\int_0^{l\omega_1/2}b(\omega)\sin(2\pi k/l\omega_1)\omega d\omega \tag{34}$$

Now, as l becomes large, let

$$2\pi/l\omega_1=\Delta T \tag{35}$$

$$(2\pi k/l\omega_1)=k\Delta T=T_k \tag{36}$$

Replace B_{ak} by $[\Delta B_a(T_k)/\Delta T]\Delta T$, and replace B_{bk} by $[\Delta B_b(T_k)/\Delta T]\Delta T$. Then, equations 33 and 34 yield 37 and 38, respectively.

$$[\Delta B_a(T_k)/\Delta T]\Delta T$$
$$=(2\Delta T/\pi)\int_0^{l\omega_1/2}a(\omega)\cos(\omega T_k)d\omega \tag{37}$$

$$\left[\frac{\Delta B_b(T_k)}{\Delta T}\right]\Delta T=(2\Delta T/\pi\times$$
$$\int_0^{l\omega_1/2}b(\omega)\sin(\omega T_k)d\omega \tag{38}$$

In the limit, as $l\rightarrow\alpha$, $T_k\rightarrow T$, hence

$$\frac{dB_a(T)}{dT}=\frac{2}{\pi}\int_0^{\infty}a(\omega)\cos(\omega T)d\omega \tag{39}$$

$$\frac{dB_b(T)}{dT}=\frac{2}{\pi}\int_0^{\infty}b(\omega)\sin(\omega T)d\omega \tag{40}$$

It is noted that as l increases K increases, but it remains true that

$$(2\pi K/l\omega_1)=K\Delta T=T_K=T_d/2 \tag{41}$$

The results expressed by equations 39 and 40 are the multipliers per unit of time T along the delay line.

It is possible, now, to establish the expressions for $A(\omega)+jB(\omega)$ with l large which is as follows

$$A_1(\omega)=\frac{\Delta T}{\pi}\int_0^{l\omega_1/2}a(\omega)d\omega+$$
$$\sum_{k=1}^{k=K}\frac{2\Delta T}{\pi}\left[\int_0^{l\omega_1/2}a(\omega)\cos\omega T_k d\omega\right]$$
$$\cos\omega T_k \tag{42}$$

In the limit l approaches infinity, T_k becomes T, and ΔT approaches zero. Then, the first term in equation 42 disappears, and

$$A_1(\omega)=\int_0^{Td/2}\left[\frac{2}{\pi}\int_0^{\infty}a(\omega)\cos(\omega T)d\omega\right]\times$$
$$\cos(\omega T)dT \tag{43}$$

It follows that

$$A(\omega)=e^{-j\omega Td/2}\int_0^{Td/2}\left[\frac{2}{\pi}\int_0^{\infty}a(\omega)\times\right.$$
$$\left.\cos(\omega T)d\omega\right]\cos(\omega T)dt \tag{44}$$

By a relatively similar process, it may be shown that

$$jB(\omega)=je^{-j\omega Td/2}\int_0^{Td/2}\left[\frac{2}{\pi}\int_0^{\infty}a(\omega)\times\right.$$
$$\left.\sin(\omega T)d\omega\right]\sin(\omega T)dT \tag{45}$$

If the upper limit of integration on the first integrals of both equations 44 and 45 became infinity as a result of increasing T_d without limit, then these expressions for $A(\omega)$ and $jB(\omega)$ would become identical with those which could have been obtained for $a(\omega)$ and $jb(\omega)$ respectively by a Fourier integral approach. Thus, the modified Fourier integral process with the time limits of 0 to $T_d/2$ yields

$$D_x(\omega)=A(\omega)+jB(\omega) \tag{46}$$

where $A(\omega)$ and $jB(\omega)$ are defined by equations 44 and 45.

TSC

Equation 23 expresses $A(\omega)$ in terms of a fixed time delay $T_d/2$ and a series of cosine terms. Following this, the time values T_1, $T_2 \ldots T_K$ were chosen such that the difference in the values of adjacent terms in the sequence was T_1. This was done because it is an essential part of the Fourier series development. However, it is conceivable that some other series of cosine terms might be formed which will give as good or better results for $A(\omega)$. In any event, suppose for the moment that no relations between T_1, $T_2 \ldots T_K$ are specified except, perhaps, that the T_k values are in the order of increasing magnitude. Then, expanding the cosine terms of equation 23 by a Taylor's series process with respect to the origin, the following results

$$A(\omega)=\epsilon^{-j\omega Td/2}\{B_{a0}/2+$$
$$B_{a1}[1-(\omega T_1)^2/2!+(\omega T_1)^4/4!-\ldots+$$
$$(-1)^{(n+\theta)/2}(\omega T_1)^{n+\theta}/K!-\ldots]+$$
$$B_{a2}[1-(\omega T_2)^2/2!+(\omega T_2)^4/4!-\ldots+$$
$$(-1)^{(n+\theta)/2}(\omega T_2)^{n+\theta}/K!-\ldots]+$$
$$\vdots$$
$$B_{aK}[1-(\omega T_K)^2/2!+(\omega T_K)^4/4!-\ldots+$$
$$(-1)^{(n+\theta)/2}(\omega T_K)^{n+\theta}/K!-\ldots]\} \tag{47}$$

where n and g are both positive real integers.

The exact expression for $a(\omega)$ given in equation 10 is a polynomial, and may be considered as its own Taylor's series expansion in which all coefficients beyond b_n are zero. If, in equation 47, K is chosen equal to $(n+g)/2$ and the successive columns of equation 47 are equated to the successive coefficients of equation 10, the following results are obtained

$$\left.\begin{aligned}
&B_{a0/2}+B_{a1}+B_{a2}+\ldots+B_{aK}=b_0\\
&B_{a1}T_1^2+B_{a2}T_2^2+\ldots+\\
&\qquad B_{aK}T_{K2}=(2!)b_2\\
&\vdots\\
&B_{a1}T_1^n+B_{a2}T_2^n+\ldots+\\
&\qquad B_{aK}T_K^n=(n!)b_n\\
&B_{a1}T_1^{n+2}+B_{a2}T_2^{n+2}+\ldots+\\
&\qquad B_{aK}T_K^{n+2}=0\\
&\vdots\\
&B_{a1}T_1^{n+g}+B_{a2}T_2^{n+g}+\ldots+\\
&\qquad B_{aK}T_K^{n+g}=0
\end{aligned}\right\} \quad (48)$$

It is possible to choose any sequence of increasing magnitudes for $T_1, T_2 \ldots T_K$, provided only that $T_K \leqq T_d/2$. Then, equations 48 may be solved for the B_{ak} values. Further, it should be noted that with this procedure the greater the value of K, the better $A(\omega)$ will approximate $a(\omega)$.

A procedure exactly analogous to that shown by equations 47 and 48, for producing an $A(\omega)$ which closely approximates $a(\omega)$, can be employed to determine $B_{b1}, B_{b2} \ldots B_{bk}$ such that $jB(\omega)$ will closely approximate $jb(\omega)$.

It should be observed that the multipliers B_{ak} obtained from equations 48 must be used at taps located T_k above and T_k below the center tap. Similar requirements are placed on the B_{bk} multipliers which would be obtained from the sine terms. If the same taps (not necessarily equally spaced) are used in determining the B_{ak} values and the B_{bk} values, then the multipliers may be combined at each tap (other than the center tap) exactly as was done for the FSC device.

References

1. Transversal Filters, H. E. Kallman, *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 28, July 1940, pp. 302–10.

2. A Phase-Plane Approach to the Compensation of Saturating Servomechanisms, Arthur M. Hopkin. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 631–39.

3. Signal Component Control, D. J. Gimpel, J. F. Calvert. *AIEE Transactions*, vol. 71 pt. II, Nov. 1952, pp. 339–43.

◆

Discussion

Abraham M. Fuchs (Westinghouse Electric Corporation, Baltimore, Md.): This paper is interesting for its results and the implications which can be drawn from the results. Figs. 4 and 6 show that the compensated systems have a linear phase lag versus frequency characteristic over a wide bandwidth. If the input signal is limited to a frequency composition within the region of linear phase lag versus frequency, then the output of the servomechanism will be an exact reproduction of the input displaced backwards in time by the reciprocal of the phase slope. As the slope is also controllable (although at the expense of higher gain in the compensator), both the accuracy of reproduction of the input signal and the time displacement of the output compared to the input can be designed into the system.

This bright side of evaluating the merits of this type of open-loop compensation is beclouded by the picture of how well said system will work when the protective coating of "linear operation" is removed. Assume, for example, that there is a load torque present (or friction on the output). The ability of the system to overcome the load torque (or friction) is not augmented by the compensator. The question arises whether the same effort spent in building the compensator might not be better spent in building the bandwidth (and generally the increased loop gain) into the control system where it not only improves the following of the input signal but also makes the control system less sensitive to load torques output friction, etc.

D. J. Ford and J. F. Calvert: We wish to thank A. M. Fuchs for his interesting discussion of our paper. As Mr. Fuchs points out, the paper deals exclusively with an open-loop type of system, and as such definitely cannot be expected to exhibit all of the performance characteristics of closed-loop systems. Assume that the system being controlled is in itself a feedback system having undesirable amplitude and phase characteristics, yet possessing the insensitivity to "nonlinearities" of the form of output friction, load torque, etc., characteristic of such feedback systems. Under these circumstances, the application of the type of compensators described in the paper will improve the over-all system response appreciably. Thus, the open-loop compensator takes the place of the elaborate correction and compensating networks in the control system which would be needed to overcome the undesirable amplitude characteristic caused by the introduction of high gain in the feedback loop.

In other words, a given system may be designed with a simple high-gain amplifier as a feedback network, thus obtaining the desirable insensitivity to output friction, load torque, etc., with little reference to amplitude and phase response. With this accomplished, the compensator may be designed to operate with this fedback control system, giving it a virtually constant gain and linear phase characteristic over as wide a frequency band as saturation limitations will permit. Thus, the combined system will possess the desirable characteristics of the closed-loop error reduction mechanism together with the desirable amplitude and phase control characteristics of the open-loop type of system described in the paper.

Magnetic Amplifiers in Metering Direct Current on Electrolytic Cell Lines

E. A. DOWNING
NONMEMBER AIEE

SEVERAL months ago every publication devoted to power and those covering industry in general carried articles telling the story of magnetic amplifiers. They told of using magnetic amplifiers with servomechanisms for automatic pilots, to control speeds of large motors, to regulate voltages of large generators, and any number of other applications, in some cases replacing equipment that had been accepted as standard for years.

The application described here is not as dramatic as this but one that is very important in the electrochemical field. That is the measurement of d-c amperes.

The particular installation has been in existence for only a matter of months but it is one of the first installations where magnetic amplifiers have been used in series to measure total d-c amperes from more than one rectifier. (Fig. 1). It is not the intention to discuss the theory of the design of magnetic amplifiers but to review the reasons for magnetic amplifiers being selected for this particular installation, to describe the method of installation, and to report on their performance.

First, since there is much similarity between magnetic amplifiers and saturable reactors, it may be well to define these as listed in the progress report of the AIEE Magnetic Amplifiers Subcommittee:

> A saturable reactor is an adjustable inductor in which the current-versus-voltage relationship is adjusted by control magnetomotive forces applied to the core.

> A magnetic amplifier is a device using saturable reactors either alone or in combination

---

Paper 54-178, recommended by the AIEE Chemical, Electrochemical and Electrothermal Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 22, 1953; made available for printing December 4, 1953.

E. A. Downing is with the Columbia-Southern Chemical Corporation, New Martinsville, W. Va.

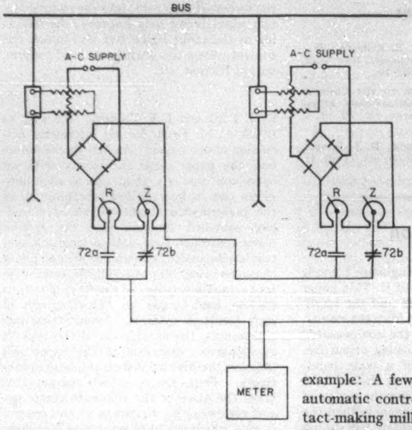


Fig. 1 (left). Schematic of current totalizing

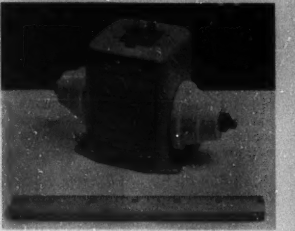


Fig. 2 (right). Simple-type magnetic amplifier

with other circuit elements to secure amplification or control.

Shown in Fig. 2 is one of the smaller magnetic amplifier units with an output rating of 40 milliwatts. This small unit has a saturable reactor with a d-c control winding on the core. The ends are connected to threaded studs to accommodate connections to shunts or other sources of control voltage. The second winding on the core has a larger number of turns, and the a-c output from this winding is rectified by a small full-wave metallic rectifier, also mounted in the case. The two windings are electrically insulated from each other.

Reasons for Using Magnetic Amplifiers

The use of magnetic amplifiers for metering current in chlorine-caustic soda cell line was elected because they would electrically isolate the recording instruments from the cell line bus bars, and because they would eliminate the cost of installing a large shunt in the main bus by totalizing the output of individual rectifiers. Any number of magnetic-amplifier units may be connected in series to totalize the largest systems.

The first-mentioned function was important because bus potential was removed from the instruments, thus eliminating a major cause of instrument failures, and removing the hazard to instrument mechanics of coming in contact with the high voltage.

In the case of certain types of instruments it is necessary that they be insulated from the higher voltages. To cite an example: A few years ago there was an automatic control application using contact-making millivolt (mv) meters. The first controller installed employed a vane and oscillator coil to detect alignment of the set pointer and indicator. As the indicator approached the set pointer it was repulsed vigorously, making it impossible to reach the control point. Other instruments failed when the insulation broke down to ground. All of these instruments were quality instruments and would have operated satisfactorily if properly used. Possibly these operating problems would have been avoided by the isolation provided by magnetic amplifiers.

The advantages of eliminating the large totalizing shunt to measure the current in the cell string were largely economical. The cost of the magnetic amplifiers at the time was less than the cost of a 50-mv totalizing shunt of adequate capacity. It turned out later that a greater saving was realized by being able to omit several feet of bus to the cell line. It had been the practice to locate all shunts in the rectifier room, for the purpose of placing them in a less corrosive atmosphere than the cell room. Elimination of the bus section to the totalizing shunt in the rectifier room saved nearly 1 ton of copper. This represents a considerable savings when the fabrication and erection cost of the copper is included. Also copper was on allocation at that time.

Another saving, which did not receive prime consideration was effected by the magnetic amplifiers. This was evident in comparing the losses in the magnetic amplifier to the losses in a totalizing shunt. The total losses of four magnetic amplifier units would be less than $1^1/_2$ watts, where the losses in a 20,000-ampere 50-mv shunt would be 1,000 watts, which amounts to more than 8,000 kilowatt-hours per year.

Installation

To install a complete set of magnetic-amplifier equipment, other accessories are required. These are a voltage regulator, an isolating transformer, a voltage step-down transformer, and a panel with potentiometers for setting the range and zero adjustments. One voltage regulator and step-down transformer can serve several magnetic-amplifier units, therefore, only one is required for each installation.

The 6-volt isolating transformers were supplied mounted on the panel with calibrating potentiometers. The other items were received and mounted separately. The magnetic-amplifier units are furnished with a set of calibrated shunt leads which were included in the initial calibration. These leads are short, making it necessary to mount the magnetic-amplifier unit near the shunt of each rectifier. In this installation the shunts are mounted behind the cathode breakers, so the magnetic-amplifier units were mounted on the frame in the rear of the circuit breaker, Fig. 3.

All of the other components may be mounted wherever it is most convenient. Their location has no effect on the performance of the equipment. All of the calibrating potentiometers were mounted in one large metal box with gasketed cover, beside the cathode breakers. The location made them convenient for testing, and the metal enclosures offers some protection from corrosive atmosphere.

Apparently the manufacturer did not consider the atmosphere around chemical plants when designing the potentiometer panel, because there was no enclosure of any kind furnished with them, Fig. 4.

Because the magnetic amplifier has to some degree the characteristics of the a-c instrument current transformer to carry a constant current regardless of the resistance in the external circuit, the potentiometers are precision-wound with manganin wire having a low temperature coefficient of resistance to eliminate errors from temperature changes.

The step-down transformer, voltage

Fig. 3. Magnetic amplifier and shunt

regulator, and recording instrument were installed in the remote panel in the station control room. The wiring between the components of this system requires very little special attention. The shunt leads must be tight and the resistance of lead connections kept as low as possible. The resistance of the complete d-c control circuit is 0.050+5=0.010 ohm. A resistance change of 0.0001 ohm would introduce a 1-per-cent error. Depending upon the installation, shielded leads might be required between the potentiometers and recorders to eliminate induced voltages in the leads.

Performance

The performance of such an installation is measured by the maintenance required and the degree of accuracy in recording current. To date the maintenance has been nil. The equipment is less than 1 year old and it may be somewhat premature to form conclusions, but there is a firm belief that it will remain very low.

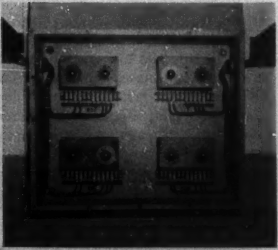


Fig. 4. Calibrating potentiometers

The subject of accuracy knits in closely with the subject of testing and calibration, so they will be discussed together. The magnetic amplifier does not replace any existing equipment but is another device added to the system. It may improve accuracy of the system by improving operation of the other instruments because maintenance requirements are decreased. A magnetic-amplifier installation is not as easily calibrated as some other installations because of the additional equipment required. The test should be made on the job site because the magnetic amplifier and associated equipment are not easily removed.

The control circuit requires approximately 5 amperes at 50 mv from the shunt at rated magnetic-amplifier input. Since the current must be variable, it is necessary to remove the leads from the shunt. A battery and variable resistor are substituted for the shunt as a d-c power supply for the control circuit. If the test is made with the magnetic-amplifier equipment in its normal position, the usual a-c power source is used without changes in connections.

Since the shunts are out of the circuits during testing, any errors introduced by the shunts are not included in the test. However, in actual operation the shunt is the source of d-c control power and therefore shunt errors are included in the measurement. Any errors introduced by the magnetic-amplifier system are added to or subtracted from the shunt error.

The published accuracy of commercial magnetic-amplifier units is 0.25 per cent, and it is understood that the tests on single units at the manufacturers' laboratory indicate a higher degree of accuracy.

The transfer or input-versus-output curve of the magnetic amplifier is not a straight line but resembles an elongated S, Fig. 5. If a straight line A-B (Fig. 5) were superimposed on the transfer curve, there is supposedly no deviation from the straight line greater than 1/4 of 1 per cent. Tests made on units in the plant show the magnetic-amplifier units well within these limits. When the straight line is extended, it is found that the indicated output is 0.58 mv with no current in the control circuit. Reportedly this 0.58-mv output with zero control current is true for all magnetic-amplifier units.

There is only one calibrating adjustment which is the range adjustment on the magnetic-amplifier equipment. This fixes the transfer curve at one point for relation of input mv. The slope of the transfer curve varies as the range is varied, but it is considered that it intersects the axis at 0.58 mv for all range settings.

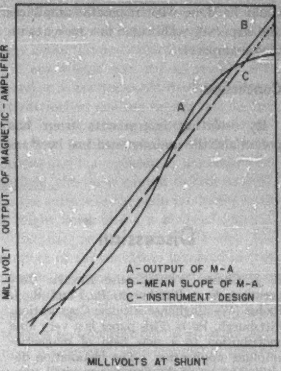


Fig. 5. Transfer curve and receiving-instrument design curve

There were minor discrepancies between field calibrations and factory tests. This was explained by stating the testing facilities at the factory were more accurate than those in our plant. No doubt the factory testing facilities are better, but an instrument to be practical must be calibrated with normal testing equipment. The field settings were accepted for final calibration.

The true accuracy of the magnetic amplifier is measured by its ability to transfer a signal from the source to a receiving instrument. This not only required near linear characteristics of the magnetic-amplifier unit, but the output characteristics of the magnetic amplifier must match the receiving instruments. Since all magnetic amplifiers have similar characteristics, the design specifications of the receiving instrument are duplicated for all installations.

Tests conducted on the equipment installed shows that the maximum over-all error is 0.3 per cent. Over a large portion of the range the error is much less than this. It is assumed that if the units were calibrated for a narrower working range, comparable to actual working conditions, the error could be reduced below 0.2 per cent.

Up to this point all discussion has been on the basis of using null balance instruments to measure the magnetic-amplifier output. As an experiment, a regular switchboard instrnment was used and the readings showed that the magnetic-amplifier output is linear with respect to the input. If one wanted to eliminate high voltage completely from the panels, it could be done with magnetic amplifiers and a specially calibrated m-v meter to indicate amperes.

Conclusion

By isolating instruments from bus potentials, the maintenance has been reduced and the average accuracy improved on the recording instruments. This has been verified here and, after only a few months' operation of one set, there were requests to install them on other units. The accuracy of the magnetic amplifier is equal to or better than most commercial instruments on the market today. For load-totalizing the initial cost of the magnetic amplifiers can usually be offset by eliminating the large totalizing shunts.

Reference

1. Magnetic-Amplifier Applications in D-C Conversion Stations, W. A. Derr, E. J. Cham. *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 220–29.

◆

Discussion

E. J. Cham (Westinghouse Electric Corporation, East Pittsburgh, Pa.) and **R. J. Radus** (Westinghouse Electric Corporation, Pittsburgh, Pa.): This paper is a very able contribution to a new field of magnetic-amplifier application. The installation described is one of the first to use small magnetic amplifiers connected to individual rectifier shunts to totalize the individual rectifier currents to obtain an indication of the total current flowing in an electrolytic cell line.

Mr. Downing has pointed out the advantages of this method of totalizing as compared to a system using a master shunt. It is of interest that this method also has advantages over a system using a single large current transductor around a totalizing bus. When all of the d-c totalizing bus passes through a current transductor, its size becomes a large factor in an installation. In addition, field calibration or field checks are impractical because of the difficulty involved in obtaining an accurate measure of the total current which usually involves thousands of amperes.

As the author pointed out, the individual small magnetic amplifiers can be checked in the field quite readily. The electrical quantities involved can be metered easily and accurately with standard instruments and equipment.

It will be well to point out that the lower portion of the transfer curve of Fig. 5 is nonlinear in excess of 1/4 per cent. This nonlinearity of operation is of no consequence in the installation described by Mr. Downing, since the rectifier units are operated at approximately full load. This means that the magnetic amplifiers are operated at all times in the linear section of their transfer curves.

Later development has produced a push-pull type of unit which exhibits a transfer curve that is linear for the full range of shunt millivolts, thus eliminating nonlinear operation for the lower values of shunt millivolts.[1] Most recent development has produced a new type of unit which combines the "to-zero" characteristic of the push-pull type of unit and the constant current characteristic of the type described in Mr. Downing's paper.

Reference

1. Industrial Applications of Transductors, R. J. Radus. *Proceedings*, National Electronics Conference, Chicago, Ill., 1953.

E. A. Downing: The discussion presented by Mr. Cham and Mr. Radus is greatly appreciated, particularly since it brings out the application of large current transductors.

I have not had the opportunity of working with transductors where tens-of-thousands of amperes are handled, but it is not difficult to imagine the problem of obtaining the large currents for calibration, or in determining accurately the value of the test currents.

The "to-zero" type of magnetic amplifier is required for some applications, but it has little advantage with chlorine cell lines other than the fact that the receiving instruments are not calibrated with a zero bias. The simple type of magnetic amplifiers are accurate and easily calibrated. In my opinion these features are preferred if it is necessary to select between these and "to-zero" characteristics.

Trends in "Automation"—Electrolytic Tinning

PAUL R. GRAVENSTRETER
MEMBER AIEE

ROBERT E. LAYTON
ASSOCIATE MEMBER AIEE

THE coating of a sheet of steel with tin is known today as tin plating. Tin plate was first manufactured by the hot-dip process started by the Welsh several hundred years ago. Early attempts were crude and largely manual. A sheet of steel of appropriate thickness and dimension was dipped in a molten bath of tin. Tinning was preceded by suitable fluxing and cleaning so as to produce a bright metallic, homogeneous coating.

In the United States, the art progressed to the point where the operations became semiautomatic. Hot dipped tin plate is made on lines for which a typical mechanical flow diagram is illustrated in Fig. 1. With the advent of the modern high-speed tandem cold-rolling strip mill, efforts were directed to the attention of a method of hot-dipping tin plate continuously. Certain lines in Europe and in this country were mildly successful, but were continuously plagued with the inability to produce a high percentage of prime sheets. Because of the setbacks that were prevalent in the continuous hot-dipping development work, electrolytic deposition of tin was considered. Little was known about the proper electrolytes, current densities and other factors which have a bearing on the end product.

To place this in time, the basic development work for the modern electrolytic tinning line was carried on in the years 1936 and 1937. The reduction in cost of producing tin plate was of paramount concern in this work. The tin plate manufacturers were constantly berated by the can manufacturers to produce tin plate for less cost. It is a matter of record that a large portion of the cost in manufacturing of any product can be traced to handling, and the damage resulting to the product therefrom. This, too, gave some push to the problem of producing electrolytic plate. By the end of 1937, a pilot line was in operation from which much data were to be gathered to permit further advancement of the art.

As a result of the occupation of the Malayan Peninsula by the Japanese, cutting off a major portion of our tin supply, tin scrap drives were instituted. The military food pack, largely in tins, forced the United States Government to impose a tin conservation program. Since the electrolytic tinning line can produce thinner coatings of tin plate than is possible by other means, electrolytic tin plating was favored by the Government. The first orders for equipment, based on the earlier experimental work, were placed in 1940.

Since then, the general acceptance and low cost of electrolytic tin plate has forced all manufacturers into this type of production. There are now approximately 37 electrolytic tinning lines in the world. Approximately 30 such units are found in the United States and Canada. In addition, the larger British companies employ three such units. On the continent of Europe there are two. Two additional units can be found elsewhere, one in Japan and another in South America.

To appreciate the size and the complexity of such an operation, refer to Figs. 2, 3, and 4. The large size of these units precludes any possibilities of manual operation. Automatic devices are considerably in evidence in the operation of this modern device. It is worth knowing that the most modern units include 11 regulating devices of various types. Each one of these regulators is a closed-loop type with internal and external types of feedback for stability.

Electrolytic tinning lines fall into three major categories depending upon the type of electrolytic solution used, namely, acid, alkali, and halogen. In addition, any one of three types of strip-heating device is used for the flow brightening of the coating.

The material produced on the original lines for the first few years was not surface-melted. That is, the tin coating was never brought up to melting temperature. To provide a better appearance, the surface of the deposited tin was brushed with rotating wire bristles. This produced a very attractive and rather bright product, but one which was not as readily soldered or as corrosion-resistant as a dense, bright melted coating or one put on the sheet by the dipping process. Some applications require a surface which is nonporous and homogeneous. To make this surface, the tin is caused to flow or melt on the surface of the strip producing both a bright finish and a nonporous coating.

Because of the complexity and the number of the machines involved in this process, automatic control or automation is required to operate the line as a unit. Speed synchronization of all sections and of each section relative to adjacent sections is required. The control of tension in the strip is also of importance within a given section. Successful strip tracking, guiding, and coil centering all depend upon constant tension. The control of storage between sections must be considered. Loops must be emptied and filled in a prescribed manner so as to afford continuity of operation. Attention must be given the problem of sheet piling in relation to the other problems. Accurate sheet spacing, overlap, consideration of defects and numerical counting must be without error. Related functions, such as plating, current density with speed changes, reflow and post treatment must each be considered to provide the proper control function at the proper time and to function with sufficient accuracy to prevent waste.

Tandem operation and its variations, e.g., shunt-down of entry section for weld, introduce complex problems of deceleration and acceleration, all of which contribute to the complexity of design. Economical co-ordinating of all these functions into an automatic operation requires sound engineering.

The number of generators to power a line of this magnitude is of concern. After analysis of the problem, two possible solutions are reached. It is possible and desirable, by certain standards, to allow one generator to power each section of the line. On the other hand, there are indications that individual generators for

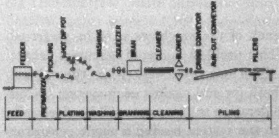


Fig. 1. Side elevation of a modern hot-dip tinning line

Paper 54-108, recommended by the AIEE Mining and Metal Industry Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing December 8, 1953.

Paul R. Gravenstreter and Robert E. Layton are with The Clark Controller Company, Cleveland, Ohio.

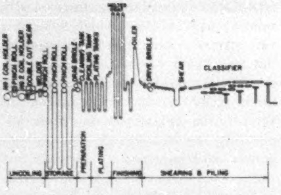

Nº 1 COIL HOLDER
PINCH ROLL
Nº 2 COIL HOLDER
DOUBLE CUT SHEAR
WELDER
PINCH ROLL
PINCH ROLL
PINCH ROLL
DRAG BRIDLE
CLEANING TANK
PICKLING TANK
PLATING TANKS
OILER
DRIVE BRIDLE
SHEAR
CLASSIFIER
UNCOILING
STORAGE
PREPARATION
PLATING
FINISHING
SHEARING & PILING

Fig. 2. Mechanical arrangement of the essential machines for an acid-type electrolytic tinning line

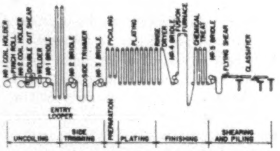

Nº 1 COIL HOLDER
PINCH ROLL
Nº 2 COIL HOLDER
DOUBLE CUT SHEAR
WELDER
Nº 1 BRIDLE
ENTRY LOOPER
Nº 2 BRIDLE
SIDE TRIMMER
Nº 3 BRIDLE
PICKLING
PLATING
RINSE
DRYER
Nº 4 BRIDLE
FUSION FURNACE
CHEMICAL TREAT
Nº 5 BRIDLE
FLYING SHEAR
CLASSIFIER
UNCOILING
SIDE TRIMMING
PREPARATION
PLATING
FINISHING
SHEARING AND PILING

Fig. 3. Mechanical arrangement of the essential machines for an alkali-type electrolytic tinning line

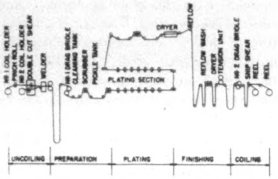

Nº 1 COIL HOLDER
PINCH ROLL
Nº 2 COIL HOLDER
DOUBLE CUT SHEAR
WELDER
Nº 1 DRAG BRIDLE
CLEANING TANK
SCRUBBER
PICKLE TANK
PLATING SECTION
DRYER
REFLOW
REFLOW WASH
DRYER
TENSION UNIT
Nº 2 DRAG BRIDLE
SNIP SHEAR
REEL
REEL
UNCOILING
PREPARATION
PLATING
FINISHING
COILING

Fig. 4. Mechanical arrangement of the essential machines for an acid- or halogen-type horizontal electrolytic tinning line

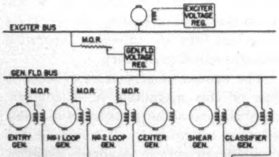

EXCITER VOLTAGE REG.
EXCITER BUS
M.O.R.
GEN. FLD. VOLTAGE REG.
GEN. FLD. BUS
M.O.R.
M.O.R.
M.O.R.
ENTRY GEN.
Nº 1 LOOP GEN.
Nº 2 LOOP GEN.
CENTER GEN.
SHEAR GEN.
CLASSIFIER GEN.
LOOP CONTROL
LOOP CONTROL
LOOP CONTROL
LOOP CONTROL

Fig. 5. Arrangement of exciter circuits for speed synchronizing of the various line sections

each of the machines is to be preferred. This problem is one of perennial debate. Much has been written on the subject, both for and against each system. Consideration in this paper will be given only to the system having a single generator-per-line section.

Two excitation systems are used to furnish generator field power. The first exciter system is the one supplying all control power for operation of contactors, mechanical brakes, auxiliary devices, and regulator reference voltages. The second furnishes synchronizing power to each of the generator main fields, Fig. 5. The tie between the general exciter and the generator field bus is by means of a voltage regulator. The drive bridle, in the processing section of the line which all other sections must match, is usually considered to be the main or "boss" drive. The voltage regulator acting on the main field supply of the center section generator establishes the general line speed by virtue of voltage, as shown in Fig. 5.

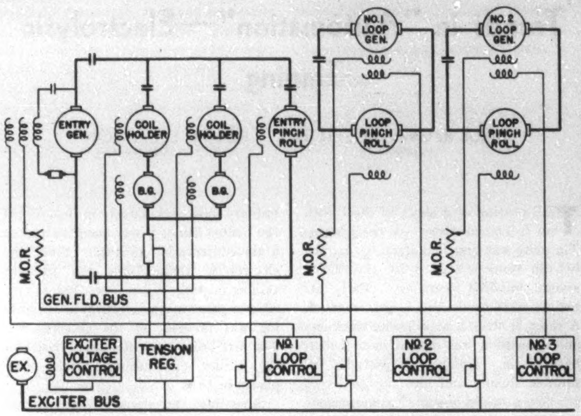

NO. 1 LOOP GEN.
NO. 2 LOOP GEN.
ENTRY GEN.
COIL HOLDER
COIL HOLDER
ENTRY PINCH ROLL
LOOP PINCH ROLL
LOOP PINCH ROLL
B.G.
B.G.
M.O.R.
M.O.R.
M.O.R.
GEN. FLD. BUS
EX.
EXCITER VOLTAGE CONTROL
TENSION REG.
Nº 1 LOOP CONTROL
Nº 2 LOOP CONTROL
Nº 3 LOOP CONTROL
EXCITER BUS

Fig. 6. Schematic diagram of the generator and regulator circuits for the entry section

Each section of the line has its own peculiarities which must be considered in providing the desired degree of automation. Consider the entry section, first as a separate device and, second, under certain conditions, as a slave to the center section. The entry section requires separate control to stop for coil change, to restart, and to synchronize with the line. In the time the entry section is down, the looping pits are emptied. On restarting they are refilled in the proper sequence so as to provide continuity of strip to the processing section. A separate adjustable voltage system is therefore supplied for the entry section. In addition, a field of the entry section main generator receives excitation from the generator synchronizing control bus (Fig. 6). Obviously, it will be necessary to overspeed the entry section to refill the loops with material used up during coil change. This is an important part of the program, as it must be accomplished with a minimum of operator attention.

The tail section of the strip, upon entering the double-cut shear, is sheared, and automatically advances to a predetermined point in the welder after shearing. A weld is made to the prepared leading end of the next coil, and the entry end restarted. The operator need not consider the operation any further, since automatic filling of the loops is provided.

Of considerable importance to the entry end operations is the proper control of tension between the entry pinch roll and the selected coil holder. This prevents looping or tearing of the material being delivered from the coil to the looping pits. Side guiding is accomplished through one of the more conventional systems.

The processing section or heart of the line, to which all other sections are referred for speed synchronization, requires particular treatment (Fig. 7). The control of tension through very narrow limits is of importance in these operations. Strip tracking through the many rolls and passes is greatly eased by the appropriate control of tension. To reflect tension as a constant value throughout the length of the processing section, helper rolls are sometimes employed in the section. These consist of either torque motors or d-c motors of low horsepower aiding in propelling the strip through the tanks.

The strip is prepared for the plating operation by first cleaning in a concentrated solution of trisodium phosphate, in some cases with the addition of electrolytic current to improve the cleaning ac-

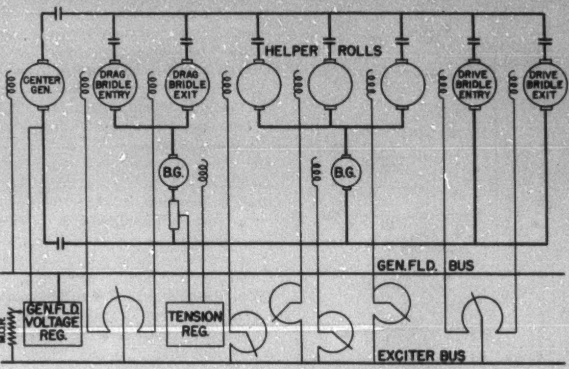


Fig. 7. Schematic diagram of the generator and regulator circuits for the center section

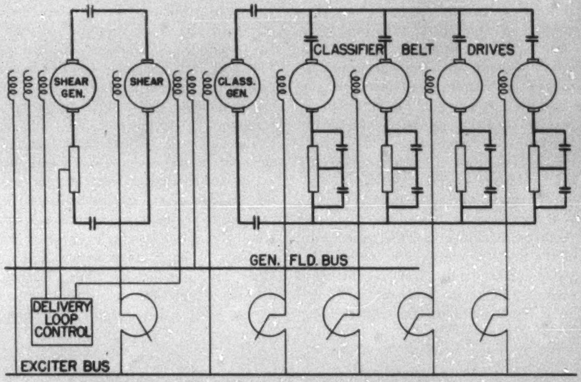


Fig. 8. Schematic diagram of the generator and regulator circuits for the delivery section with armature acceleration for the classifier belt conveyors

tion. Cleaning is a process employed to remove oil and grease picked up in prior processing. To remove the scale as occasioned by annealing, the strip is pickled in an electrolytic bath of sulfuric acid. Plating of tin on the strip of steel is accomplished in one of three methods. In the first method an acid electrolyte consisting of stannous sulfate may be used. In the second method a solution of alkaline electrolyte, consisting principally of sodium stannate with sodium hydroxide additions, is employed. The third method is with a solution containing tin chlorides and fluorides, and is known as the halogen electrolyte method. The mechanical arrangement in the plating tanks for each of these solutions differ widely. In some operations, where wide line speed variations are to be experienced, current density is maintained by a change in the number of passes used for plating. The acid system requires that the current be held in proportion to strip speed and regulation of the plating current is therefore mandatory.

The alkaline electrolyte, on the other hand, has a very narrow current density range requiring very accurate speed maintenance. As a result of the accurate speed, no current regulation is required once all of the variables reach their steady-state performance level.

The alkaline line, instead of having accurate current regulation, requires accurate speed regulation. The halogen line is not widely used, and data on the operation of this system are very meager. Current densities of each of the solutions, at approximately the same electrolyte temperature, are as follows: Acid, up to 400 amperes per square foot; alkaline, 20 to 30 amperes per square foot; halogen line, approximately 185 amperes per square foot. Cathodic and anodic efficiencies in both the acid and halogen process are reportedly 100 per cent, while the alkaline bath is in the neighborhood of 70 per cent. Since the work function of the metal is involved, the energy to transfer the metal tin to the strip in any of these processes is essentially a constant.

Three general methods are in use for accomplishing the reflow part of the process. Listed in the order of their use, the means of obtaining the necessary heat are resistive heating, radiation heating, and induction heating. Resistive heating consists of passing a current through the strip to produce heat. Radiation heating consists of either an electric or gas-fired furnace surrounding the strip. In the third method, induction heating produces heat in the surface of the strip using a high-frequency (100 kilocycles) type of furnace or chamber.

Because certain oxides or stains form on the surface of the sheet under storage, post treatment of the tin plate is necessary. At this writing, there is considerable variance in the proposed methods of post treatment. The process is still under development with considerable further study indicated. Certain types of tin plate require oiling. This may be accomplished by either of two methods: first, by oiling the strip through suitable rolls designed for the purpose; and, second, by electrostatic means which charges atomized particles of oil electrically and deposits them on the strip in the way dust is collected in an electrostatic dust precipitator.

The usual delivery section of the line consists of one of two operations. The strip may be coiled or it may be sheared into sheets of appropriate size which are then classified according to defects. These defects range from surface defects to pin holes and they are detected by the electronic pin hole detectors and by flying micrometers. After automatic classification, approximately 1,100 prime sheets are counted off into a package.

Successful piling is related to the sheet spacing after shearing. The sheet spacing should remain a constant, regardless of speed, above a certain threshold. During shearing, the delivery section does not stop with the rest of the line but continues to run at the minimum speed so

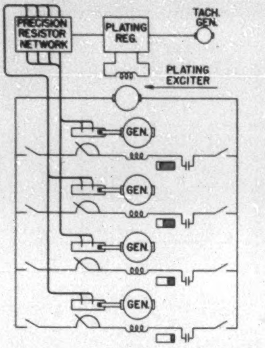


Fig. 9. General connections of the plating current regulator

as to clear the conveyors of sheets and avoid cobbling. This runoff is usually timed automatically and is not under the control of the operator. Also, while shearing, it is necessary to maintain a loop ahead of the shear so that the measuring rolls of the shear are always faced with the same drag of material (Fig. 8). This is accomplished through the use of a modulated loop control which accurately determines the position of the loop and modifies the delivery section generator voltage control to control the loop.

These processes are required to run in synchronism and under the supervision of the line operator, which involves the use of many tandem relaying circuits, each of which compels its section of the lines to respond through the synchronizing ties. The entire line may be started and stopped by the operation of a single push button. Individual group control stations are provided along the line to assist the operators in threading, maintenance, or other noncontinuous operations. The real intelligence and sense for line operation comes from the master control station. Here the line speed is selected, the plating current adjusted to produce the desired coating, and the melted coating current control adjusted to produce the right surface flow conditions. All other associated factors are also controlled from the master control station. Circuit design is required which provides continuity of operation, but also permits the operator to make decisions which will resolve difficulties that arise. Lines of this type may operate continuously without a shutdown for something like 160 hours, with 8 hours down for general maintenance, to cover a week's operation.

Various regulator circuits have been advanced over a period of years to fulfill the requirements of tension, speed, current, etc., each of which is a large enough subject for a paper, in itself. Block diagrams indicate the complexity of such regulators. A high degree of accuracy is required of each of these units, being of the order of 0.5 per cent (Figs. 9 and 10).

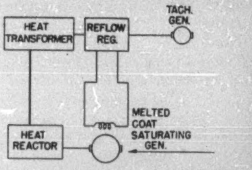


Fig. 10. Arrangement of the most important components forming the melted coat regulating circuit

Conclusions reached, after careful consideration of all the factors, indicate the problem to be one of many ramifications. No one problem is solved without considering the implications with regard to other processes or functions. Careful analysis of each relay or control function forces co-ordination with all other functions. The general solution to the problem can be found in isolating or sectionalizing line functions into groups and attention devoted to these problems, after which combinations and recombinations occur, and a compatible scheme, may be reached. Automation, a new large-scale problem to the control designer, multiplies the engineering work to a point where cost of engineering is considered a factor of sizable proportions compared to the equipment supplied. Automation, an outgrowth of a search for lower production costs and improved products, has increased the complexity beyond the scope of the untrained engineer. Indications tend to stress the need for automation specialists within the control industry.

References

1. MULTIPLE GENERATORS FOR PROCESSING LINES, **E. E. Vonada.** *Iron and Steel Engineer*, Association of Iron & Steel Engineers, Pittsburgh, Pa., vol. 29, Nov. 1952, pp. 77–86.

2. DIFFERENTIAL THICKNESS COATING, **S. S. Johnson.** *Iron and Steel Engineer Year Book*, Pittsburgh, Pa., 1952, pp. 280–83.

◆

Discussion

C. C. Thomas (General Electric Company, Schenectady, N. Y.): The authors have written an interesting and informative paper and I agree with their refusal to be drawn into a discussion on the relative merits of individual versus common generators for tinning lines. It is primarily a matter of economics but each case should be considered separately, weighing such factors as cost, duplication of equipment, simplicity of circuits, off standard generator ratings, etc., before a decision is made.

In discussing trends in automation of electrolytic tinning lines, however, I do believe the authors should have pointed out that a very definite trend is toward differential coating of the strip. This causes additional complications in the switching, metering, and regulating circuits of the coating generators. The object of this, of course, is to save tin by coating only the inside of the tin can with a coat heavy enough to withstand the acids, etc. The outside needs only a light coat so the can will not rust on the grocers' shelves in a reasonable period of time. Coating the tin plate heavy on one side and light on the other incidentally brings up interesting problems of marking it so each side can be readily identified by the fabricator. At present one of the most common methods is to use a sandblasted roll on one side in the final temper pass given before the strip is brought to the tinning line. This gives one side of the strip a dull finish even after coating. Usually this dull side is coated the heaviest and, therefore, goes on the inside of the can where appearance is not a problem. Other good schemes for marking would be welcomed by the industry.

The trend toward constantly increasing speeds of tinning lines has created problems which make us take a look at torque-regulated helper roll drives. Many of these helper roll drives are submerged in liquid and at these higher speeds the roll apparently pumps liquid between the strip and the roll. Because of this liquid which serves as a lubricant, torque-regulated helper roll motors only aggravate the speed mismatch between them. Since the speed of the helper rolls is no longer tied to the strip by friction it is necessary to lay out a control which will economically maintain approximate speed correspondence at all times at the higher speed and still not push or pull too hard on the strip at the lower speeds when the roll and strip are locked together.

And finally, mention should be made that the highest speed tinning lines do not have the flying shear and classifier section on the delivery end. Unless radical improvements are made in the shear and classifier the trend will be to build more lines without these units and do the shearing and classifying as a separate operation.

Paul R. Gravenstreter and **Robert E. Layton**: We wish to thank C. C. Thomas for his interest in discussing this paper.

In a paper of this type, it is difficult to establish what material is to be included and what omitted as relatively unimportant. The point of view, when writing the paper, is of importance in selecting the material to be included. Differential coating of tin plate falls into this classification. While being important as related to tin saving, it is not considered as being important from the standpoint of automation. Automatic plating current control is not primarily a problem involving connection of plating generators, for differential plating, as a problem of total plating current and its relation to line speed.

Mr. Thomas' reference to the proper control for helper roll drives is very timely. Electrolytic tin plate producers for some years have been working with the problem of eliminating roll marking of the strip due to slippage. Ideas have been advanced, tried, and proved unsatisfactory in the past. Speed-regulated helper rolls are now under consideration as Mr. Thomas has indicated. This system can become rather expensive if many groups, or individual rolls, are subject to regulation. Time alone will indicate the success of this latest attempt to eliminate the problem of idler roll strip marking.

Future line designs will be dictated, to a large degree, by economics. Coil handling and damage, as a result of handling, is an expensive item in the tin plate industry. The problem of handling tin plated coils to be edge trimmed and sheared is another step in the manufacture of tin plate when compared with a shearing and classifying type of line. Tin plate, produced at very high speeds, which is coiled after plating, involves the use of a number of expensive low-voltage plating generators, particularly when heavy coating weights are produced. Problems of this type are resolved by the user, who must consider these and other factors, such as royalties on patented processes, etc. These problems are normally not placed before the drive designer, yet greatly influence the general line design to which the drive must be adaptable.

A Graphical Procedure for Determining the Gain of a Servomechanism for a Specified Maximum Modulus Less Than Unity

THOMAS J. HIGGINS
MEMBER AIEE

A PROBLEM of frequent occurrence in servomechanism analysis and design is that of determining the gain K of the forward-frequency transfer function

$$[G(s)]_{s=j\omega}=\left[\frac{K(1+a_1s+\ldots+a_ms^m)}{s^k(1+b_1s+\ldots+b_ns^n)}\right]_{s=j\omega} \tag{1}$$

of a single-loop unity feedback system, Fig. 1, such that the over-all frequency transfer function

$$\frac{C(j\omega)}{R(j\omega)}=\frac{G(j\omega)}{1+G(j\omega)}=\frac{|G(j\omega)|e^{j\alpha_1}}{|1+G(j\omega)|e^{j\alpha_2}}=Me^{j(\alpha_1-\alpha_2)}=Me^{j\alpha} \tag{2}$$

has a specified maximum modulus $M=M_m$. Accordingly, those textbooks[1-4] which present a comprehensive integrated account of basic servomechanism theory advance the outline of a simple graphical procedure for determining the required gain K, providing $M_m>1$.

Often, however, it happens that the designer desires the value K which yields a specified value of $M_m<1$. For example, in the oft-encountered case of Fig. 2, wherein the feedback transfer function $H(s)$, though nonunity, is yet a pure numeric, $H(s)=K_h$, we have, from well-known theory

$$\frac{C(j\omega)}{R(j\omega)}=\frac{G(j\omega)}{1+G(j\omega)H(j\omega)}=\frac{1}{H(j\omega)}\times\frac{G(j\omega)H(j\omega)}{1+G(j\omega)H(j\omega)} \tag{3}$$

and thus

$$\frac{C(j\omega)}{R(j\omega)}=\frac{1}{K_h}\,\frac{K_hG(j\omega)}{1+K_hG(j\omega)} \tag{4}$$

Defining $G_1(j\omega)$ by

$$G_1(j\omega)=K_hG(j\omega) \tag{5}$$

and substituting accordingly in equation 4 yields

$$\frac{C(j\omega)}{R(j\omega)}=\frac{1}{K_h}\,\frac{G_1(j\omega)}{1+G_1(j\omega)}=Me^{j\alpha} \tag{6}$$

Now the second equality in equation 6 can be written in the equivalent form

$$\frac{G_1(j\omega)}{1+G_1(j\omega)}=K_hMe^{j\alpha}=M_1e^{j\alpha} \tag{7}$$

wherein

$$M_1=K_hM \tag{8}$$

In virtue of equation 8, it follows that the determination of the gain K of $G(j\omega)$, such that M in equations 2 and 6 takes on a stated maximum value M_m can be phrased as: the determination of the gain K_1 of $G_1(j\omega)$ such that $M_{1m}=K_hM_m$. Now, in practice, M_m is usually specified as $M_m=1.4$ or so, and often* K_h is a fractional value, say $K_h\cong 0.25$. In such case, $M_{1m}=K_hM_m\cong 0.35$; and thus the designer is faced with precisely the above mentioned problem: namely, the determination of gain $K_1=KK_h$ of the frequency transfer function $G_1(j\omega)$ such that the maximum value of the modulus M_1 of the "equivalent" over-all frequency transfer function delineated by equation 7 is of a specified value $M_{1m}<1$.

However, although all the aforementioned servomechanism textbooks advance a simple graphical procedure for determining the gain for a specified maximum modulus M_m greater than unity, none advance such a procedure for the just obtained case of maximum modulus M_m less than unity; nor is such to be found in the periodical literature. Obviously, however, it is most desirable that such procedure be available to the servo analyst and designer. Accordingly, the prime purpose of this paper is to fill this gap through advance of the desired graphical procedure.

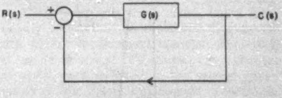


Fig. 1. Block diagram of single-loop unity feedback system

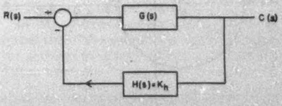


Fig. 2. Block diagram of single-loop nonunity feedback system

---

* For example, in the use of feedback through a tachometer or a voltage divider, as evidenced in *An Illustrative Example*.

Paper 54-127, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 16, 1953; made available for printing November 30, 1953.

Thomas J. Higgins is with the University of Wisconsin, Madison, Wis.

The derivation of the graphical construction in the direct plane stemmed, in part, from discussion with William Kippenhan, a former graduate student in my servomechanism class.

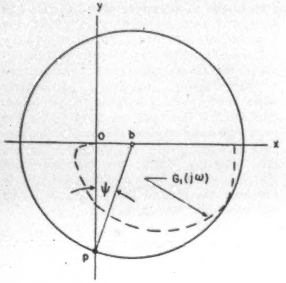


Fig. 3. Basic construction in the direct plane

Graphical Procedure in the Direct Plane

From well-known theory, the radius and the center (x_o, y_o) of a circle of constant value of $M_1<1$, located in the direct (x, y) plane defined by $G_1(j\omega)=x+jy$, are

$$\text{Radius}=M_1/(1-M_1^2);\ (x_o, y_o) \equiv [M_1^2/(1-M_1^2);\ 0] \quad (9)$$

Consider, now, a typical circle for $M_1<1$, as indicated in Fig. 3. Let (pb) be a line segment directed perpendicular to the tangent at p, where p, as indicated, is the point of intersection of the negative y-axis and the circle of constant M_1. Then, perforce, the directed segment (pb) passes through the center b of the circle. Accordingly, it follows from equation 9 that

$$(pb)=\text{radius}=M_1/(1-M_1^2);\quad (ob)=x_o=M_1^2/(1-M_1^2) \quad (10)$$

and thus

$$\psi=\sin^{-1}[(ob)/(pb)]=\sin^{-1}M_1 \quad (11)$$

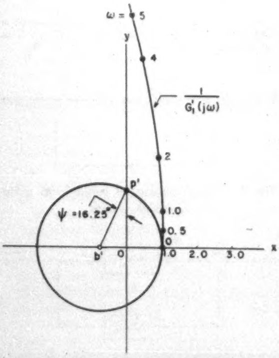


Fig. 4. Basic construction in the inverse plane

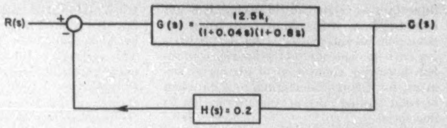


Fig. 5 (right). Block diagram of illustrative system

Let, now, M_1 be the specified maximum value $M_1=M_{1m}$. Then the plot of the frequency transfer function

$$G_1(j\omega)=\left[\frac{K_1(1+a_1's+\ldots+a_m's^{m'})}{s^{k'}(1+b_1's+\ldots+b_{n'}'s^{n'})}\right]_{s=j\omega} \quad (12)$$

whereof K_1 is the desired value, must be tangent, as in Fig. 3, to the circle of radius and center given by equation 10, wherein $M_1=M_{1m}$.

In turn, these considerations yield the desired graphical construction for K_1 to be as follows:

1. Plot $G_1'(j\omega)$, obtained from equation 12, by taking $K_1=1$ therein.
2. With center b' on the positive x-axis, strike arcs tangent to the plot of $G_1'(j\omega)$ until a point p', the intersection of the negative y-axis and the arc, is located such that the line segment $(p'b')$ connecting the point p' and the center b' makes an angle $\psi=\sin^{-1}M_{1m}$ with the y-axis.
3. Measure the distance (ob'). Then increase of the scale of graphical construction by a factor K_1, such that $K_1(ob')=(ob)$ as in equation 10, would convert the plot of $G_1'(j\omega)$ of the originally scaled construction into the plot of $G_1(j\omega)$; and as on the rescaled construction the circle is both tangent to $G_1(j\omega)$ and has the radius and center of equation 10, whereof $M_1=M_{1m}$, it follows that the scale factor K_1 is the desired value of K_1. Accordingly, $K_1=(ob)/(ob')$; or, substituting for (ob) from equation 10

$$K_1=M_{1m}^2/[(1-M_{1m}^2)(ob')] \quad (13)$$

wherein M_{1m} is as specified and (ob') is as measured from the graphical construction typified by Fig. 3, assuming primed quantities thereon.

Obviously, with the value of K_1 in hand, the actually desired value K of the forward transfer function $G(j\omega)$ is, by equation 5, given by $K=K_1/K_h$.

An easily effected alternative for step 2 above is to incline one edge of a draftsmen's triangle through the origin o at an angle $\psi=\sin^{-1}M_{1m}$ with the y-axis; and holding this angle constant, to then displace the triangle downwards until the distance $(p'b')$ and center b' thus laid off are such that the arc of radius $(p'b')$ and center b' is tangent to the plot of $G_1'(j\omega)$.

Graphical Procedure in the Inverse Plane

A corresponding graphical procedure can be easily deduced for use in the inverse (x, y) plane defined by $x+jy=1/G_1(j\omega)$. In this case the radius and the center (x_o, y_o) of a circle of constant value of M_1, are, by well-known theory

$$\text{Radius}=1/M_1;\ (x_o, y_o)\equiv(-1, 0) \quad (14)$$

Consider, now, a typical circle for $M_1<1$, as indicated in Fig. 4. Let (pb) be a line segment directed perpendicular to the tangent at p, where p, as indicated, is the point of intersection of the positive y-axis and the circle of constant M_1. Then, perforce, the directed segment (pb) passes through the center b of the circle. Accordingly, it follows from equation 14 that

$$(pb)=\text{radius}=1/M_1;\ (ob)=|x_o|=1 \quad (15)$$

and thus

$$\psi=\sin^{-1}[(ob)/(pb)]=\sin^{-1}M_1 \quad (16)$$

Let now M_1 be the specified maximum value $M_1=M_{1m}$. Then the plot of the inverse of the frequency transfer function $1/G_1(j\omega)$, where $G_1(j\omega)$ is as defined by equation 12, must be tangent, as in Fig. 4 to the circle of radius and center given by equation 15, wherein $M_1=M_{1m}$.

In turn, these considerations yield the desired graphical construction for K_1 to be as follows:

1. Plot $1/G_1'(j\omega)$, obtained from equation 12, by taking $K_1=1$ therein.
2. With center b' on the negative x-axis, strike arcs tangent to the plot of $1/G_1'(j\omega)$ until a point p', the intersection of the positive y-axis and the arc, is located such that the line segment $(p'b')$ connecting the point p' and the center b' makes an angle $\psi=\sin^{-1}M_{1m}$ with the y-axis.
3. Measure the distance (ob'). Then increase of the scale of the graphical construction by a factor $1/K_1$, such that $(ob')(1/K)=1$ as in equation 15, would convert the plot of $1/G_1(j\omega)$ of the originally scaled construction into the plot of $1/G_1(j\omega)$; and as on the rescaled construction the circle is both tangent to $1/G_1(j\omega)$ and has the radius and center per equation 15, whereof $M_1=M_{1m}$, it follows that the scale factor $1/K_1=1/(ob')$ encompasses the desired value of K_1. Accordingly

$$K_1=(ob') \quad (17)$$

wherein (ob') is as measured from the graphical construction typified by Fig. 4.

As in the foregoing, with the value of K_1 in hand, the actually desired value K of the forward transfer function $G(j\omega)$ is, by equation 5, given by $K=K_1/K_h$.

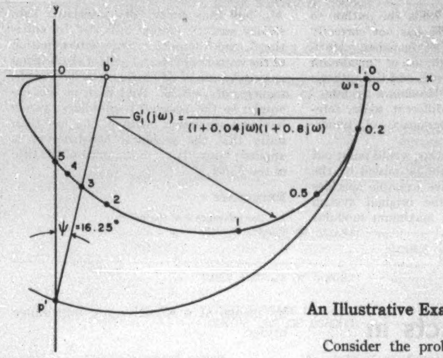


Fig. 6. Construction in the direct plane for the illustrative system

Again, an easily effected alternative for step 2 is to incline one edge of a draftsman's triangle through the origin o at an angle $\psi=\sin^{-1}M_{1m}$ with the y-axis; and holding this angle constant, to then displace the triangle upwards until the distance $(p'b')$ and center b' thus laid off are such that the arc of radius $(p'b')$ and center b' is tangent to the plot of $1/G'(j\omega)$.

An Illustrative Example

Consider the problem of determining the parameter k_1 of $G(s)$ of the servo system of Fig. 5 such that $M_m=1.4$ (which problem is no. 9.8 of p. 476 of reference 1). Then, by equation 5 and the data of Fig. 5

$$G_1(s)=0.2G(s)=2.5k_1/(1+0.04s)(1+0.8s) \quad (18)$$

The plot of $G_1'(j\omega)=1/(1+0.04j\omega)(1+0.8j\omega)$ is as shown in Fig. 6. Striking arcs with centers on the positve x-axis until a center b' is located, such that the line segment $(p'b')$ is inclined to the y-axis at, by equation 11, an angle

$$\psi=\sin^{-1}M_{1m}=\sin^{-1}M_mK_h=\sin^{-1}0.28=16.25° \quad (19)$$

yields $(ob')=0.22$. Hence, by equation 13, the gain K_1 of $G_1(j\omega)$ is

$$K_1=(0.28)^2/[1-(0.28)^2]0.22=0.386 \quad (20)$$

Accordingly, from equations 18 and 20, we have $2.5k_1=0.386$; and thus $k_1=0.155$. Corroboratively, computation of M_m from equation 3 using this value of k_1 in $G(j\omega)$ yields $M_m=1.4$.

References

1. SERVOMECHANISMS AND REGULATING SYSTEM DESIGN (book), H. Chestnut, R. W. Mayer. John Wiley and Sons, Inc., New York, N. Y., 1951, pp. 233–36, 242–43.

2. PRINCIPLES OF SERVOMECHANISMS (book), G. S. Brown, D. P. Campbell. John Wiley and Sons, Inc., New York, N. Y., 1948, pp. 185–89.

3. SERVOMECHANISM ANALYSIS (book), G. J. Thaler, R. G. Brown. McGraw-Hill Book Company, Inc., New York, N. Y., 1953, pp. 184–90.

4. PRINCIPLES OF AUTOMATIC CONTROLS (book), F. E. Nixon. Prentice-Hall, Inc., New York, N. Y., 1953, pp. 145–47.

◆

Discussion

E. J. Vredenburgh (University of Missouri, Columbia, Mo.): This paper fills a clear need of a simple graphical procedure for determining the gain of a servomechanism with a specified maximum modulus less than unity. The alternative construction suggested seems to be the most practical to use and compares favorably with construction methods given in texts for the case of maximum modulus greater than unity.

Frank Petree (Phillips Petroleum Company, Idaho Falls, Idaho): The author asserts that it often happens that a designer is faced with the problem of determining a value of open-loop gain for a feedback controller, such that the peak value of the modulus M is less than unity. It supports this argument by reference to a specific problem concerning a type-O regulator, which the author claims to have a feedback constant which is less than unity; this is erroneous, as we shall soon see.

It appears that the author was misled because the problem in question did not specifically define what was meant by "a maximum of 1.4 for its steady-state frequency response." In the problem, the feedback element is a tachometer whose transfer constant was given as 0.2 volt-seconds per radian. It is evident that, had the problem chosen to use other units, the feedback constant might have been 2.39 volt per rpm, for instance. In this case, the forward gain of the controller would, of course, undergo a similar alteration, from 12.5 radian per volt-second to 1.31 rpm per volt. Thus we see that care must be taken in the particular case where the reference does not possess the same dimensions as the controlled variable. A mature designer will carefully define the controlled variable in terms of the reference, so as not to be led to a ridiculous solution. For instance, the author arrives at a controller whose open-loop gain is 0.39 and whose closed-loop gain at zero frequency is 1.395 volt-second per radian. If we define the maximum modulus as the ratio of the maximum gain, 1.4, to the zero-frequency gain, 1.395, it certainly does not appear that the criterion of $M_m=1.4$ is met!

A realistic way of attacking the problem is to describe the controlled variable in units of the reference as determined by the zero-frequency response of the feedback element. This is equivalent to using the feedback quantity, rather than the controlled variable, in the determination of M_m.

Alternatively, we can describe the controlled variable in units of the reference as determined by the zero-frequency closed loop response of the regulator. This is, perhaps, a more precise method, but is nearly equivalent to the first if the open loop gain is large, or if the controller is other than type-O.

One might argue that controllers are often designed in which the feedback constant is a numeric and is still less than unity. This does not alter the problem, for one must still scale the output to units of the input, or a completely misleading result will obtain.

Thomas J. Higgins: The author would like to express his sincere appreciation of Professor Vredenburgh's remark that "this paper fills a clear need of a simple graphical procedure for determining the gain of a servomechanism with a specified maximum modulus less than unity"—which remark by a well-known and experienced teacher provides substantiation of the author's own reason for advancing the paper.

Next, the author concurs with Professor Vredenburgh that the alternative construction is most practical (i.e., easiest to carry out), and is on a parallel re ease of use and degree of accuracy with the well-known graphical construction method for the case of maximum modulus M_m greater than unity.

In Mr. Petree's first paragraph it is stated that ". . . the author claims to have a feedback constant which is less than unity; this is erroneous. . . ." To the contrary, however, the author makes no such claim, nor is he in error. Rather, not the author, but Messrs. Harold Chestnut and Robert W. Mayer,[1] authors of the book from which the illustrative problem is cited, advance the problem—and it is not a claim, but a fact, clearly manifested on the block diagram of Fig. 5, that the feedback constant $H(s)=0.2$ is less than unity! Of course, by a change of units, such as cited in the second paragraph of Mr. Petree's discussion, a feedback constant greater than unity can be obtained, but this has no bearing on the fact that when working in terms of the dimensional units in which the solution is actually carried out, the feedback constant is less than unity.

In Mr. Petree's second paragraph it is stated that "the author was misled because the problem in question did not specifically define what was meant by a maximum of 1.4 for its steady-state frequency response." However, the author would remark he was not misled, and that specific definition of a maximum M_m of 1.4 is given; see immediately after equation 2, where M_m is defined as the maximum value of the over-all (or closed-loop) frequency transfer function. This symbolism, and corresponding definition, are standard by any of the authoritative texts on servomechanism theory, a great body of published periodical literature, and, I believe, the recommendations of the Subcommittee on Terminology of the Feedback Control Systems Committee. In such fact, Mr. Petree's last remark in his second paragraph, "If we define the maximum modulus as the ratio of maximum gain, 1.4, to the zero-frequency gain, 1.395, it certainly does not appear that the criterion of $M_m = 1.4$ is met!" leads the author to believe that Mr. Petree has not correctly assimilated the content of the paper. However this may be, though, use of "maximum modulus" in his own sense, and the author's clearly defined use of "maximum modulus" in a well-established different sense, forestalls mutual correspondence of understanding.

In conclusion, the author would point out that a reader should not be misled by the fact that the illustrative example specifies $M_m = 1.4$. Although the original system thus does not have a maximum modulus M_m less than unity, the associated subsidiary analytic system indicated by equation 7, and from which stems determination of the numerical value of gain of the original system by use of M-circles, does have such: namely, $M_{1m} = 0.28$. And it is, in this example, to the associated subsidiary system having a maximum modulus M_{1m} less than unity that the graphical construction is applied; correctly so, in the spirit of the title of the paper.

REFERENCE

1. See reference 1 of the paper.

Iron Conduit Impedance Effects in Ground Circuit Systems

A. J. BISSON
MEMBER AIEE

E. A. ROCHAU
NONMEMBER AIEE

THE resistance and reactance of a circuit increases when associated with an iron conduit. In normal circuit operation where the current-carrying conductors are wholly within a conduit, the iron has little effect on the resistance and reactance since the magnetic field locally surrounding the conductors is partly in iron and mostly in air. However, the circuits encountered in building wiring during a phase-to-ground fault condition that involves the conduit in and around the faulted circuit are subject to an appreciable increase in resistance and reactance. The purpose of this paper is to record a relatively simple method of approximating the impedance of such a faulted circuit, to point out the significance of the impedance values as they occur in practice and why these values must be considered by the engineer in design of circuits, particularly under faulted conditions.

The results of the data show that, for the circuit arrangement selected, the determination of the circuit a-c resistance and reactance can be accomplished through the use of impedance characteristics of a given length of the iron conduit itself.

Impedance Determination

The calculation of the impedance of a single-phase circuit in iron conduit is laborious and inaccurate, and not convenient for use by practicing engineers. When iron becomes a part of the electric circuit, increases in both resistance and reactance of the circuit will be large. Moreover, the resistance and reactance of the circuit will vary, depending upon the amount of current being carried.

To avoid the difficulties of such calculations, a simple method based on the results of a series of laboratory tests is presented from which to determine, in a practical way and to a reasonable degree of accuracy, circuit impedances for three single-phase circuit arrangements. These circuit arrangements are frequently found during phase-to-ground faults in building wiring.

The impedances of the particular circuit arrangements in question have been found to be simple numerical multiples of the impedance of the iron conduit itself. The multiples are determined from test data for the following three conditions:

Case 1. Current through iron conduit with remote return circuit, 10-foot spacing.

Case 2. Current through cable in iron conduit with remote return circuit, 10-foot spacing.

Case 3. Current through cable within a conduit, the latter forming the return circuit.

The laboratory tests were made on conduit runs 50 feet in length using 3-inch and 4-inch iron pipe and 500,000-circular-mil (CM) cable. Measurements of 60-cycle resistance and reactance were made by means of a sensitive a-c impedance bridge.

Results of Test

An analysis of the test results shows the following:

1. For a remote return circuit, the impedance of the current-carrying cable in the conduit is twice the basic impedance of the corresponding length of iron conduit (case 2).
2. The impedance of the iron conduit may be considered equal to zero when single-phase current in a conductor within the duct is returned on the duct itself (case 3).
3. When single-phase current in a conductor within the iron conduit is returned on the conduit itself, all losses are reflected on the cable whose impedance is increased to be approximately equal to the basic impedance of the conduit itself (case 3).

As an explanation of the test results, a simple analogy is presented in Fig. 1, showing by means of schematic diagrams the approximate current distribution in the iron conduit wall. The distribution of current with respect to the conduit wall surfaces is as follows:

Case 1—Current in outer surface only.

Case 2—Current in both inner and outer surfaces.

Case 3—Current in inner surface only.

In accordance with accepted a-c theory, current in iron conductors penetrates to only a relatively small depth, the penetration depending upon the physical dimensions, resistivity, and magnetic characteristics of the iron.

There are a number of methods by which this penetration may be calculated.[1-3] From calculated values and confirming test results, it is concluded that the skin effect is high, resulting in a-c to d-c resistance ratios of 10 to 1 and 12 to 1 respectively, for the 3-inch and 4-inch conduit at a current of 250 amperes. These ratios are apparent from the test results which show the d-c and a-c resistances respectively for 1,000 feet of 3-

---

Paper 54-36, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 16, 1953; made available for printing November 17, 1953.

A. J. BISSON and E. A. ROCHAU are with the Consolidated Edison Company, New York, N. Y.

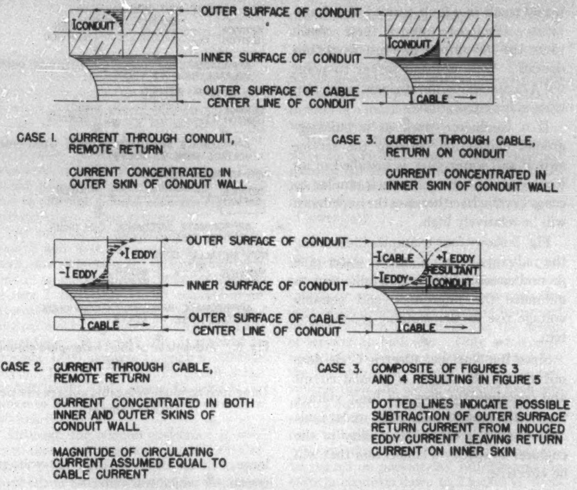

I CONDUIT
OUTER SURFACE OF CONDUIT
INNER SURFACE OF CONDUIT
OUTER SURFACE OF CABLE
CENTER LINE OF CONDUIT
I CONDUIT
I CABLE
CASE 1. CURRENT THROUGH CONDUIT, REMOTE RETURN
CURRENT CONCENTRATED IN OUTER SKIN OF CONDUIT WALL
CASE 3. CURRENT THROUGH CABLE, RETURN ON CONDUIT
CURRENT CONCENTRATED IN INNER SKIN OF CONDUIT WALL
+I EDDY
-I EDDY
I CABLE
OUTER SURFACE OF CONDUIT
INNER SURFACE OF CONDUIT
OUTER SURFACE OF CABLE
CENTER LINE OF CONDUIT
-I CABLE
+I EDDY
RESULTANT
-I EDDY = I CONDUIT
I CABLE
CASE 2. CURRENT THROUGH CABLE, REMOTE RETURN
CURRENT CONCENTRATED IN BOTH INNER AND OUTER SKINS OF CONDUIT WALL
MAGNITUDE OF CIRCULATING CURRENT ASSUMED EQUAL TO CABLE CURRENT
CASE 3. COMPOSITE OF FIGURES 3 AND 4 RESULTING IN FIGURE 5
DOTTED LINES INDICATE POSSIBLE SUBTRACTION OF OUTER SURFACE RETURN CURRENT FROM INDUCED EDDY CURRENT LEAVING RETURN CURRENT ON INNER SKIN

Fig. 1. Probable current distribution in wall of conduit

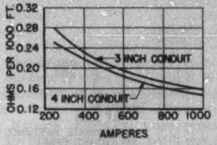


Fig. 2. Sixty-cycle resistance of iron conduit at 25 degrees centigrade

D-c resistance: 3-inch conduit, 0.0288 ohm per 1,000 feet; 4-inch conduit, 0.0192 ohm per 1,000 feet.

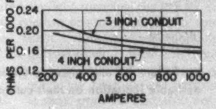
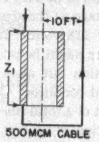
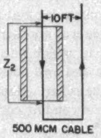


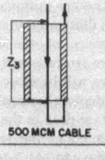
Fig. 3. Sixty-cycle reactance of iron conduit at 25 degrees centigrade

Table I. Sixty-Cycle Circuit Impedance in Ohms per 1,000 Feet

Amperes	3-Inch Conduit			4-Inch Conduit		
Case 1: Impedance of iron conduit with remote return circuit						
	R_1	X_1	Z_1	R_1	X_1	Z_1
250	0.280	0.224	0.360	0.252	0.194	0.318
500	0.212	0.186	0.282	0.200	0.176	0.266
1,000	0.160	0.164	0.229	0.148	0.154	0.214
Case 2: Impedance of cable in iron conduit with remote return circuit						
	R_2	X_2	Z_2	R_2	X_2	Z_2
250	0.540	0.436	0.694	0.514	0.380	0.639
500	0.440	0.380	0.581	0.416	0.348	0.542
1,000	0.340	0.320	0.467	0.330	0.310	0.453
	R_2/R_1	X_2/X_1	Z_2/Z_1	R_2/R_1	X_2/X_1	Z_2/Z_1
250	1.93	1.95	1.93	2.04	1.96	2.01
500	2.08	2.04	2.06	2.08	1.98	2.04
1,000	2.12	1.95	2.04	2.23	2.01	2.12
Case 3: Impedance of cable with iron conduit as return circuit						
	R_3	X_3	Z_3	R_3	X_3	Z_3
250	0.280	0.212	0.351	0.254	0.190	0.317
500	0.234	0.192	0.308	0.210	0.176	0.274
1,000	0.188	0.160	0.247	0.180	0.146	0.232
	R_3/R_1	X_3/X_1	Z_3/Z_1	R_3/R_1	X_3/X_1	Z_3/Z_1
250	1.00	0.95	0.98	1.01	0.98	1.00
500	1.10	1.03	1.07	1.05	1.00	1.03
1,000	1.18	0.98	1.08	1.22	0.95	1.08


10FT
Z_1
500 MCM CABLE


10FT
Z_2
500 MCM CABLE


Z_3
500 MCM CABLE

D-c resistance of 500,000-CM cable is 0.0216 ohm per 1,000 feet.

inch iron conduit to be 0.0288 and 0.28 ohm, and for the 4-inch conduit 0.0192 and 0.25 ohm. Effects on cable associated with such conduits will therefore be of the same order of magnitude.

The current distribution indicated in Fig. 1 has been drawn to approximate that obtained by calculation. The horizontal lines of these current areas are used to illustrate the relative magnitude of current at various depths of penetration.

The analogy shown fits the test data for the three cases in every respect, the losses being distributed as follows (see Fig. 1):

Case 1—Single area, losses in outer skin only.

Case 2—Two areas, losses in both inner and outer skin.

Case 3—Single area, losses in inner skin only.

No attempt has been made to separate the iron losses into their hysteresis and eddy current components, which by a combination of physical laws appear to adjust themselves so as to produce the results presented herein.

From results 1 and 3 in the foregoing, the impedance of these single-phase circuits is largely a function of the impedance of the iron conduit. A comparison of the resistance component with the d-c resistance of the cable (Table I) indicates that, for the conductor sizes normally used in 3-inch and 4-inch conduit, the size of the cable conductor has relatively little effect on the circuit impedance. This statement has been confirmed by tests with first one, then four and lastly seven 4/0 conductors in parallel in 4-inch conduit at 200 amperes which showed impedance values of 0.72, 0.66 and 0.66 ohm respectively per 1,000 feet.

The a-c resistance and reactance characteristics of the 3-inch and 4-inch iron conduits (case 1) are plotted in Figs. 2 and 3 as a function of current. Characteristics such as these, which are quite easy to obtain on any specimen of conduits, may be used as a basis for determining single-phase circuit impedance.

In Table I is given a set of representative values of resistance, reactance, and impedance for the three cases, together

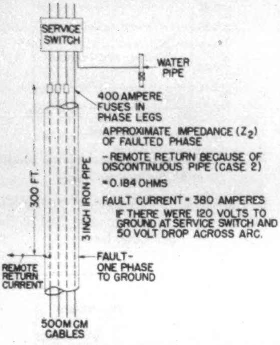

SERVICE SWITCH
WATER PIPE
400 AMPERE FUSES IN PHASE LEGS
APPROXIMATE IMPEDANCE (Z2) OF FAULTED PHASE
- REMOTE RETURN BECAUSE OF DISCONTINUOUS PIPE (CASE 2)
= 0.184 OHMS
FAULT CURRENT = 380 AMPERES
IF THERE WERE 120 VOLTS TO GROUND AT SERVICE SWITCH AND 50 VOLT DROP ACROSS ARC.
300 FT.
3 INCH IRON PIPE
FAULT - ONE PHASE TO GROUND
REMOTE RETURN CURRENT
500M CM CABLES

Fig. 4. Failure of one phase-to-ground undesirable limitation on fault current

with computed ratios showing the relation existing between cable impedance and the basic impedance of the iron conduit itself.

In case 3 of Table I the complete circuit consists of the conductor and return over the conduit, but only impedance values of the cable are given because the 60-cycle potential along the conduit was practically zero. All the voltage drop in the circuit appeared along the cable within the conduit. The reason for this apparent paradox of current in the pipe but no voltage drop is essentially that the voltage drop resulting from the current in the conduit is offset by a voltage rise induced in the conduit by the current in the conductor.

Some Practical Applications

When one phase leg of a 3-phase circuit is grounded to an iron conduit, the path of the short-circuit current is similar to case 3 if the conduit is continuous. This would be so, even though there may be a parallel path external to the conduit. If the conduit is not continuous, then case 2 applies and the impedance is correspondingly higher. The resulting high impedances may be such that the circuit fuse will not operate.

The following example of such a condition is illustrated in Fig. 4. From Figs. 2 and 3, the resistance and reactance of a 3-inch iron pipe at 400 amperes is 0.235 and 0.195 ohm per 1,000 feet respectively. For 300 feet, the impedance value is 0.0915 ohm. For case 2, the cable impedance is twice the pipe impedance and equals 0.183 ohm. A ground from one conductor to the pipe at 300 feet as shown, if there were a 50-volt drop across the arc and 120 volts at the service switch, would result in a fault current of 380 amperes. Obviously, under these conditions, the 400-ampere line fuse would not operate. This problem serves to point out a factor which the engineer must consider in his circuit design.

If a conductor, such as a building-ground connection between the service switch and water pipe, is installed in an iron conduit, the condition is similar to cases 1 or 2. In either case the impedance will be relatively high.

Fig. 5 shows four examples illustrating the advantage of a short water-pipe ground connection in nonmetallic pipe to minimize the impedance and possible voltage rise in the building-ground system.

Since the National Electric Code does not specify the length of ground circuit and consequently does not limit voltage rise on the ground connection under fault condition, it is the responsibility of the engineer to specify a connection that will be effective.

Although the effect of phase isolation is well known, the magnitude of the effect in unusual situations may not be generally realized. As an example, a customer complained of low voltage on one phase of a 3-phase 4-wire underground service. On test, the voltage at the service switch was found to be 115, 115, and 101 volts. The voltage at the street mains was 117, 117, and 118 volts. Initially it was thought there might be poor contact at some connection but all joints were cool and in good condition. It was found, however, that of a total run of 130 feet, there was a section of 35 feet where two phase legs were in one iron conduit and the remaining phase leg and the neutral in another. There was another section of mains of 60 feet in which one phase leg was in one pipe and two phase legs and the neutral in the other. Replacement of the cables in one conduit corrected the unbalanced condition.

Another example of an unusual situation illustrates an unforeseen effect of the presence of the iron conduit. When there are two conductors in an iron pipe, there will be close coupling between the conductors because of the iron pipe. The voltage drop in the first conductor will be reflected in the second. The action is similar to a 1-to-1 transformer. In an actual instance of a single-phase impedance test on a long section of circuit there were two conductors in an iron pipe, one a high-voltage circuit carrying test current and the other presumably a low-voltage circuit with negligible current. As outlined in case 2, the impedance of the circuit was high, resulting in a large voltage drop in the high-voltage conductor which was reflected in the low-voltage circuit with undesirable results.

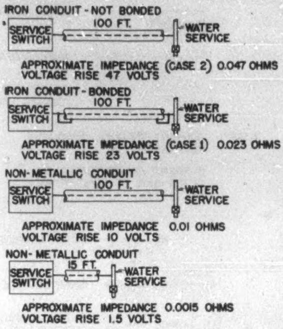

IRON CONDUIT - NOT BONDED
100 FT.
SERVICE SWITCH
WATER SERVICE
APPROXIMATE IMPEDANCE (CASE 2) 0.047 OHMS
VOLTAGE RISE 47 VOLTS
IRON CONDUIT - BONDED
100 FT.
SERVICE SWITCH
WATER SERVICE
APPROXIMATE IMPEDANCE (CASE 1) 0.023 OHMS
VOLTAGE RISE 23 VOLTS
NON-METALLIC CONDUIT
100 FT.
SERVICE SWITCH
WATER SERVICE
APPROXIMATE IMPEDANCE 0.01 OHMS
VOLTAGE RISE 10 VOLTS
NON-METALLIC CONDUIT
15 FT.
SERVICE SWITCH
WATER SERVICE
APPROXIMATE IMPEDANCE 0.0015 OHMS
VOLTAGE RISE 1.5 VOLTS

Fig. 5. Advantage of short water-pipe ground in nonmetallic pipe

Three-inch conduit, 4/0-cable voltage rise per 1,000 amperes

From 3, under "Results of Test," it follows that in case of a conductor grounded to the pipe of a high-voltage pipe-type cable, all losses are reflected on the cable and there is little likelihood of any appreciable 60-cycle rise in pipe potential with respect to ground. There is little known, however, of the effects of transient currents returning along the pipe. This problem needs further investigation.

Conclusions

Use of the method outlined affords a simple and reasonably accurate means of determining single-phase circuit impedances in an iron conduit. The method is particularly useful for checking the effectiveness of fuse protection in building wiring under fault conditions. The advantages of short ground-wire runs in nonmetallic conduit are proved, and indicate the need for continued serious consideration by the engineer. The simple analogies provided are an aid in understanding some of the actions of single-phase circuits in iron conduit.

References

1. ALTERNATING CURRENT CONDUCTIVITY OF IRON PIPE, **H. B. Dwight.** *The Electric Journal*, East Pittsburgh, Pa., vol. 23, 1926, p. 295.

2. THEORY AND CALCULATION OF TRANSIENT ELECTRIC PHENOMENA AND OSCILLATIONS, **C. P. Steinmetz.** 3rd ed, pp. 383–385.

3. EXPERIMENTAL RESEARCHES ON THE SKIN EFFECT OF STEEL RAILS, **A. E. Kenelly, F. H. Achard, A. S. Dana.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 182, Aug. 1916, pp. 135–89.

Discussion

R. H. Kaufmann (General Electric Company, Schenectady, N. Y.): The significance of this paper reaches far beyond simply the impedance effects in ground circuits. It lays the groundwork for understanding the factors responsible for peculiar current distribution patterns in ground circuits associated with a-c system operation. This should stimulate further investigation and lead to new and revised concepts regarding equipment grounding systems.

The contrast in test results obtained in cases 2 and 3 is most striking. In both cases, current is conveyed through the 3-inch conduit on a 500,000-CM copper conductor. In case 2, the return circuit consists of an external 500,000-CM copper conductor running parallel to the conduit with about a 10-foot spacing. About 47 volts is required to circulate 1,000 amperes through a 100-foot length. In case 3, the return circuit is the conduit itself. Only about 25 volts is required to circulate 1,000 amperes through a 100-foot length.

Although the conduit resistance is over seven times that of an external 500,000-CM copper cable, a ground return circuit formed by the conduit requires only a little more than half as much voltage as would be required for the external copper cable return circuit. It is the difference in circuit reactance which is responsible for this result. This clearly indicates the strong tendency for the ground return current to seek a path in close proximity to the outgoing conductor in spite of a much lower resistance in a more remote circuit conductor.

May this stimulating introductory paper serve to spearhead other investigations expanding our knowledge of ground circuit behavior. This knowledge in turn will lead to more effective protective ground systems of lower cost.

E. J. Rutan (Chemical Construction Corporation, New York, N. Y.): The authors of this paper are making available to engineers data which influence the design of feeder and branch circuits in buildings and the design of services to buildings. Only a few engineers have been aware of the importance of the impedance due to steel conduits in preventing fuses and circuit breakers from protecting circuits during faults. The authors' examples are somewhat limited, but these data will permit extended analysis.

From tests I have made faults may not be cleared on present-day 120/208-volt circuits in conduits down to 2 inches in diameter and 200 feet in length. These hazards are not existent in the higher voltage (440-volt) circuits. These hazards are aggravated by the fact that present fuse and circuit-breaker standards permit clearing times at 200 per cent of rated current of 6 to 12 minutes in ratings of 100 amperes and above. I brought these conditions to the attention of the National Electrical Code revision committee in December 1952, but it was too late for consideration in the new code.

These effects also enter into service connection design. In the National Electric Safety Code as early as 1944 recognition of some hazards due to these effects on 208-volt alternating network services were provided for.

To briefly touch on some of the methods by which these potential hazards may be eliminated I might mention the use of nonmagnetic conduit or mineral insulated metal sheathed cable. Bare neutral conductors in steel conduit will help. Fuse and circuit-breaker standards should be changed and maximum operating time at low overloads reduced. It may even be desirable not to ground the service neutral.

Codes and Standards are developed from the collective experience of various engineers. The application of the data in this paper should result in the suggestion of many revisions as engineers analyze circuit and service installations and faults.

Rectifier Motive Power—Inductive Co-ordination Considerations

E. B. KING
MEMBER AIEE

K. H. GORDON
MEMBER AIEE

L. J. HIBBARD
MEMBER AIEE

FOR many years, the use of rectifier motive power was considered impracticable for several reasons, one of which involved the telephone interference problem. Corrective measures considered necessary to minimize power supply wave distortion of the single-phase rectifier to a degree which would be satisfactory for telephone communication rendered this type of motive power unacceptable to locomotive builders and to railroads.

Advances in the rectifier and communication arts since the early studies have tended to reduce substantially some of the deterrents to the use of rectifiers for motive power. From the rectifier standpoint, the development of successful large-capacity sealed-type rectifiers and the consequent elimination of the vacuum pump difficulties were important advances.

From the communication standpoint, the extensive replacement by railroad and telephone companies of open wire with cable, with its inherent shielding and close pair spacing, as well as the increasing use of carrier, have made the communication plant less susceptive. Another factor has been the general policy of telephone companies for many years to avoid long and close inductive exposures to existing or potential railroad electrifications.

It is therefore apparent that many factors have favored the present application of rectifier locomotives to existing electrifications. These factors would not be present to the same extent along routes involving lighter trackage and which have not heretofore been envisioned as subject to electrification. With proper design and co-ordination, however, no difficulty from telephone interference would be expected on such routes.

Early Operating Experience

An experimental rectifier-type multiple-unit (MU) car has been in continuous service with originally applied apparatus on The Pennsylvania Railroad since July 14, 1949. Except for test purposes, no corrective apparatus has been connected on this equipment. The car which has excellent power factor and efficiency draws 52 amperes from the overhead 11-kv catenary during acceleration. No added or different noise in adjacent communication circuits has been attributed to operation of this first rectifier-type motive power unit.

Tests were made on the experimental car on the Philadelphia–West Chester branch in August 1949 and near Wilmington, Del., in September 1949. A 2-car train was used, consisting of the rectifier car no. 4561 and a-c motor car no. 764. Either car could be used singly to supply the propulsion with the other car acting as a load.

The results of these tests are summarized as follows:

1. The influence of the rectifier car, without filters, as measured by the $I \cdot T$[*] product

---

Paper 54-110, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing November 20, 1953.

E. B. KING is with the American Telephone and Telegraph Company, New York, N. Y., K. H. GORDON is with The Pennsylvania Railroad, Philadelphia, Pa., and L. J. HIBBARD is with Westinghouse Electric Corporation, East Pittsburgh, Pa.

[*] $I \cdot T$ represents the current in amperes supplied to the car times the telephone influence factor which has a different weighting for each frequency in the voice range, based upon telephone receiver response.

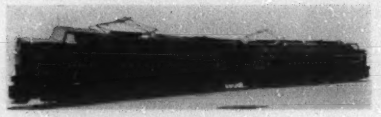


Fig. 1 (left). Two-unit 6,000-hp single-phase rectifier locomotive

was of somewhat lower magnitude than that resulting from the a-c motor car.

2. The use of filters of different types on the rectifier car gave very large reductions (up to a maximum of about 80 per cent) in the $I \cdot T$ product.

3. The prominent rectifier harmonics were below 425 cycles, but for the a-c motor car, the principal harmonics were in the order of 2,000 cycles, likely resulting from the slot harmonics of the motors. There were also sizable low-frequency harmonics in the order of 75-175 cycles probably resulting from the saturating effects from the iron in the motors.

4. The maximum $I \cdot T$ products of the trolley currents measured on various equipments were:

Rectifier car without filter	1,500
Rectifier car with best filter	300
A-c motor car without filter	2,000
A-c motor car with best filter	900
P5a freight locomotive (a-c motors)	1,400
GG1 passenger locomotive (a-c motors)	2,250

Two 6,000-horsepower (hp), rectifier-type locomotives have operated on The Pennsylvania Railroad without corrective apparatus connected except for test purposes. Locomotive 4995-4996, Fig. 1, entered revenue service in November 1951; locomotive 4997-4998 in February 1952. These 2-unit locomotives normally draw 750 to 800 amperes from the 11-kv overhead wires, with occasional short-time peaks of 1,000 amperes.

Tests were made on locomotive 4995-4996 during the summer and fall of 1951 on main-line tracks between West Yard and Davis, south of Wilmington, Del. On these tests, one unit hauled the dynamic braking load of the other unit. Railroad telephone personnel, listening on adjacent telephone circuits during these tests, could detect no added noise directly chargeable to rectifier motive power. Similar results were obtained when locomotive 4997-4998 was tested between West Yard and Davis.

After locomotive 4995-4996 entered revenue service, further listening tests on railroad telephone circuits made by different operating personnel were reported to have detected no added noise for several weeks. However, when using loud speakers on some of these railroad circuits, power directors at Harrisburg noticed a distinctive noise when a rectifier locomotive was operating in certain portions of their territory. The characteristic noise produced by rectifier motive power will be described in detail later.

Factors Involved in Co-ordination Problem

Inasmuch as the theory of induction between power or electrified railroad circuits and neighboring communication circuits has been adequately covered in previous publications, no attempt will be made to include a complete discussion in this paper. It may be well to recall, however, that the over-all effect of a given inductive exposure depends upon three factors: 1. inductive influence of the power system; 2. inductive coupling between the power system and the telephone system; 3. inductive susceptiveness of the telephone system.

If any of these three factors could be reduced to zero, the inductive effects would disappear. In the present state of the art, it is generally known that it is impracticable to reduce any of these factors to zero. The coupling is usually fixed by the practical routes of the facilities involved. Therefore, in most cases the solution to inductive problems involves keeping both susceptiveness and influence at reasonably low values.

Inductive influence includes those characteristics of a supply circuit, in this case the catenary system with its associated apparatus, which determine the character and intensity of the inductive field which it produces. The susceptiveness of the communication circuits is a function of the type and degree of balance of these circuits. The control of telephone susceptiveness is largely a matter of controlling unbalances and the control of power system influence is almost entirely a matter of controlling harmonics.

Telephone circuits which pass through a-c magnetic fields will have induced voltages which may produce currents in connected telephone receivers. These extraneous currents will interfere with telephone transmission, if they are in the audio or noise frequency range, and are of appreciable magnitude compared with normal voice currents.

The voltage induced in telephone circuits can be resolved into two components: metallic circuit induction (unequal voltages will be induced in the two conductors if they are located in different strengths of an a-c field) and longitudinal circuit induction. If the two conductors of a telephone circuit are subjected to the same average field strengths by means of transpositions, the voltages induced in the two conductors by an a-c field will be equal.

Longitudinal voltages impressed on a perfectly balanced circuit will cause no current in bridged telephone receivers. However, when either series or shunt unbalances are present, longitudinal induced voltages acting upon them will produce noise frequency currents in the telephone receivers.

In telephone circuits entirely in grounded sheath cable, induced noise results solely from magnetically induced longitudinal-circuit voltages and currents acting on telephone circuit unbalances. The electric shielding effect of the grounded cable sheath eliminates electric induction, and the continuous twist and close spacing of the pairs prevent any appreciable magnetic induction directly in the metallic circuit. In telephone cables with open-wire extensions exposed beyond the end of the cable, metallic-circuit induction in the open wire may be important, and this induction is not of course affected by cable sheathing.

Since it is impracticable to maintain perfect balance in any communication circuit, induced voltages will be present and there is some value of longitudinal in-

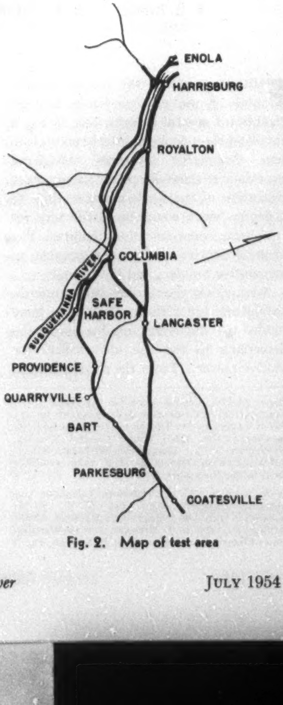


Fig. 2. Map of test area

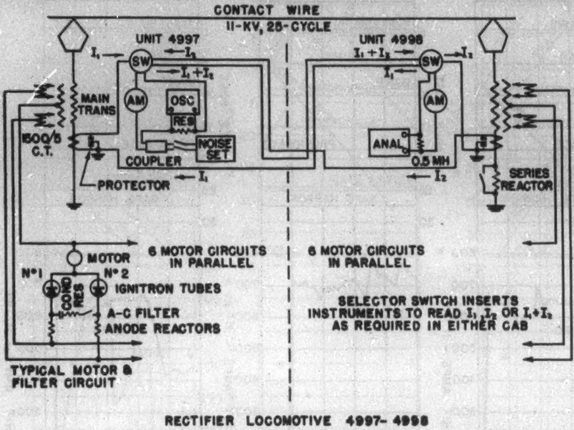

CONTACT WIRE
11-KV, 25-CYCLE
UNIT 4997
UNIT 4998
I_1
I_2
I_1+I_2
SW
AM
OSC
RES
NOISE SET
MAIN TRANS
1500/5 C.T.
COUPLER
PROTECTOR
ANAL
0.5 MH
SERIES REACTOR
6 MOTOR CIRCUITS IN PARALLEL
6 MOTOR CIRCUITS IN PARALLEL
MOTOR
N°1
N°2
CAPS
RES
IGNITRON TUBES
A-C FILTER
ANODE REACTORS
TYPICAL MOTOR & FILTER CIRCUIT
SELECTOR SWITCH INSERTS INSTRUMENTS TO READ I_1, I_2 OR I_1+I_2 AS REQUIRED IN EITHER CAB
RECTIFIER LOCOMOTIVE 4997-4998

Fig. 3. Connections for inductive interference tests

duced voltage which, if exceeded, may cause undue impairment of telephone transmission efficiency.

It is generally well known that speech involves a range of frequencies from about 125 cycles to about 10,000 cycles but that satisfactory communication can be achieved with a considerably narrower band of frequencies, and telephone circuits usually transmit such narrower bands. If extraneous voltages or currents induced on telephone circuits, as just described, are to cause impairment of ability to converse over the telephone, they must be within the frequency range transmitted by the telephone circuit and reproduced by the telephone receiver. Obviously the wider the frequency range handled by the telephone system, the wider the range within which extraneous currents and voltages can cause noise.

Although the trend toward wider frequency bands, which bring with them improvements in the ease of telephone communication, has, in a sense, made the noise co-ordination problem more difficult, there have been concurrent developments in the design, construction, and maintenance of the telephone lines and central office and station equipment which have reduced to a substantial degree the effects arising out of unbalance. For example, whereas metallic circuits of 25 to 30 years ago were 50 to 200 times better from the induction standpoint than a ground return circuit, the best modern circuits may be 1,000 times or more better. Similar improvements have taken place in central office circuits and subscriber telephone sets

Co-ordinated Noise and Influence Tests

The first series of tests to determine the effect of rectifier motive power on neighboring communication circuits was made with experimental rectifier-type (MU) car 4561 during the summer of 1949. Three additional series of tests were made on rectifier-type electric freight locomotives during April and May of 1952, October of 1952, and April of 1953.

Table I. Summary of Road Tests

Run No.	Filters	Locomotive $I \cdot T$	Noise-to-Ground,* Decibels (Db)	Metallic Noise, Db
First Series				
1	None	24,000	51	34
2	Reactor	12,000	44	32
3	Ten capacitors per motor	3,400	39	18
4	Ten capacitors per motor and reactor	2,200	35	19
Second Series				
5	None		53	36
6	None, one *P5a* locomotive		36	24
Third Series				
7	Five capacitors per motor	4,000	37	26
8	Three capacitors per motor	6,500	40	27
9	One capacitor per motor	14,500	48	32

* Noise-to-ground on a telephone circuit is also a measure of influence, and bears almost a direct relationship with the $I \cdot T$ product measured on the locomotive.

Tests on MU car 4561 showed clearly that telephone noise effects produced by the rectifier-type propulsion apparatus were well below the values which had been accepted with the conventional a-c MU equipment operated by The Pennsylvania Railroad. Although no corrective measures have been applied to this car, except during these tests, the nature of mitigative apparatus which might be economically used on larger power units was determined.

All of the rectifier locomotive units were initially equipped with a-c filters, and a dephasing reactor was installed for test purposes on locomotive unit 4998. Separate filter circuits are provided for each of the six traction motors on each locomotive unit. The filter circuit consists of one adjustable damping resistor connected in series with a variable number of capacitors. Ten 23.5-microfarad 1,800-volt capacitors, each rated 25 kva on 60 cycles or about 11 kva on 25 cycles, and weighing approximately 75 pounds, are available in each filter circuit. The dephasing reactor, mounted on unit 4998, weighs 1,450 pounds and is designed for 1.83 ohms at 382 amperes. Taps are arranged at approximately 10, 40, 60, and 80 per cent of 1.83 ohms.

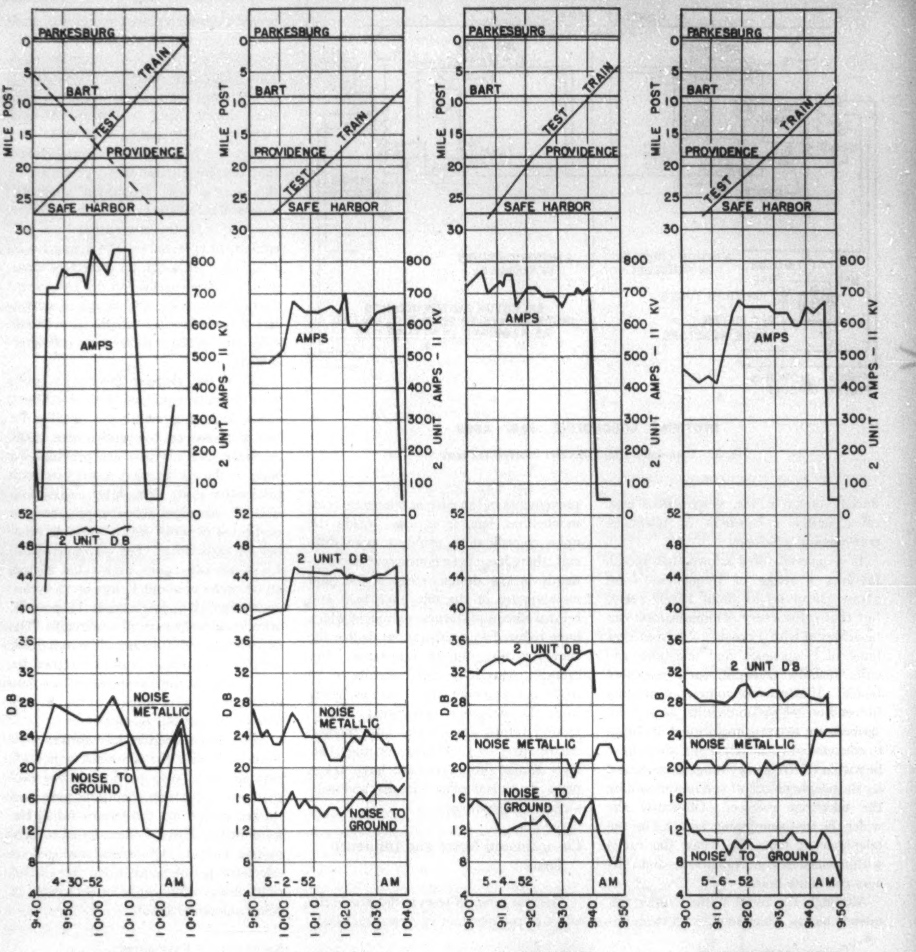
Tests on the rectifier locomotives were conducted with the locomotives hauling revenue trains over portions of the railroad where exposure conditions were considered promising. Observers riding the locomotives collected data related to locomotive output. Observers stationed at selected exposure points along the railroad right-of-way obtained records of noise in communication circuits.

Selection of Exposures

First reports of a "different" noise which was chargeable directly to rectifier motive power came from power direc-

Table II. Noise-to-Ground Data

Type of Locomotive	Load, Flat Tons	Maximum Noise-to-Ground, Db	Remarks
None, minimum reading		24	Other sources
One *P5a*	8,459	33	
Two *P5a*	6,752	36	
Two *GG1*	4,277	33	
E2B (2-unit)	7,550	34	



NO FILTER
NO REACTOR

REACTOR ONLY

10 CAPACITORS
PER MOTOR

10 CAPACITORS/MOTOR
WITH REACTOR

Fig. 4. Data obtained from first series of locomotive tests

tors whose duties required frequent use of a railroad telephone circuit between Harrisburg and Parkesburg, Pa., Fig. 2. This circuit, 68 miles in length, was contained in tape-armored aerial cable, paralleled by a 4/0 copper aerial neutralizing wire separated about 12 inches for shielding, and mounted on the railroad right-of-way in close proximity to the electrified freight tracks over which the rectifier locomotive operated. Twenty-two miles of the circuit were exposed to the 0.30 per cent ruling grade of the freight tracks between Columbia and Parkesburg where full capacity of the locomotives was utilized.

Also exposed to full load of the locomotive on the long ruling grade was a telephone company toll circuit from Quarryville to Lancaster, Pa., a distance of about 15 miles. The open-wire section of this circuit crossed the railroad at Quarryville and separation from the railroad was highly irregular throughout its length. Earth resistivity in the region is high. The railroad at the nearest part of the exposure is on a high fill. The central office and its entrance cable are so close to the tracks that there is probably an appreciable ground potential effect when electric locomotives pass Quarryville. Transposition layout of the open-wire section was made on the basis of neutraliz-

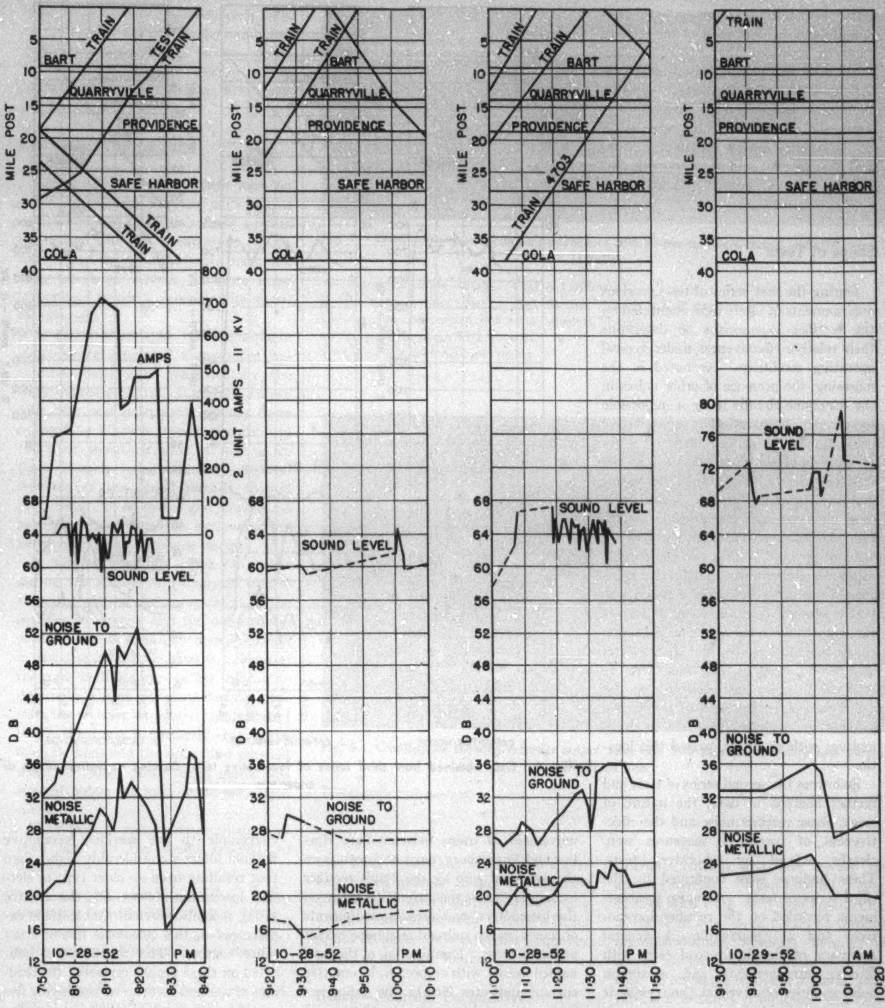


Fig. 5. Data obtained from second series of locomotive tests

ing induction effect from adjacent 60-cycle power lines rather than for neutralizing the effect of the 25-cycle railroad exposure. The exposue is so irregular with respect to the railroad that it is probably impracticable to attempt to obtain improved transposition layout with the 25-cycle railroad circuit.

Other exposure locations selected initially included subscribers' lines at Columbia, Pa. and Langhorne, Pa., both operated by The Bell Telephone Company of Pennsylvania, and open-wire or aerial cable circuits at Trenton, Lawrence, Deans, and Outcalt, N. J., all operated by the New Jersey Bell Telephone Company. The rectifier locomotive produced only small increases in noise on these exposures, and observations at these locations were discontinued early during the initial series of locomotive tests.

Measuring Instruments

Noise readings were recorded on test locomotives by a harmonic analyzer and a noise meter, arranged as shown in Fig. 3, so that 11-kv ampere readings and $I \cdot T$ noise influence readings could be taken on either locomotive unit or on both units simultaneously. At most of the selected exposure locations, the only test instrument used was a Western Electric type *2B* noise meter. A Western Electric harmonic analyzer and two tape recorders were used in addition to the noise meter at Parkesburg during the tests of April and May of 1952, and at Quarryville during the tests of October 1952 and April 1953.

Scope of Tests

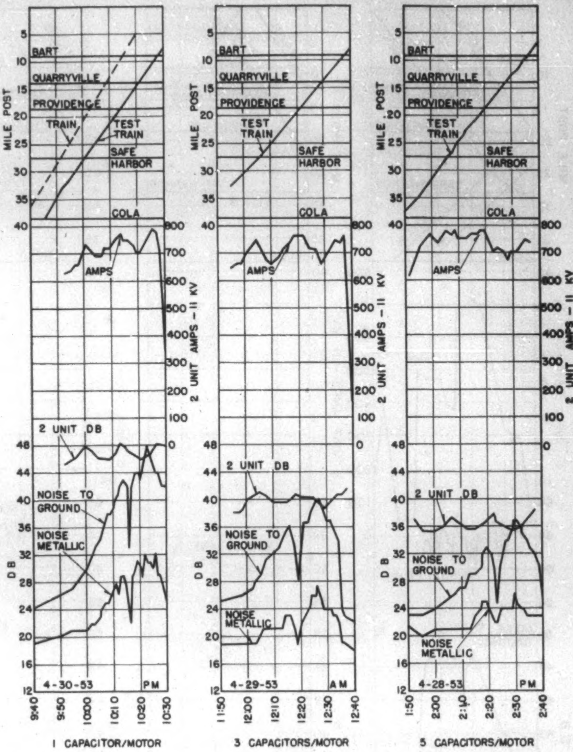
During the first series of tests, various arrangements of filters were connected on the rectifier locomotives to determine their relative effectiveness under typical operating conditions. As noted in the following, the presence of other noises in the telephone circuits made it impossible to develop a practicable quantitative evaluation of the various filters.

Analysis of the data compiled in the first tests clearly indicated: 1. that noises produced by older types of electric locomotives occasionally exceeded in magnitude the noise produced by the rectifier locomotive; and 2. that the Quarryville exposure, while subject to inductive influences from other industrial sources, was better adapted than the other exposures for quantitative evaluation of induction chargeable to the rectifier locomotive. The second series of tests was therefore confined to noise measurements on the Quarryville exposure as various electric locomotives, including the rectifier locomotives with no filters, passed this location.

Following the second series of tests and further analysis of data, the nature of single-phase rectifier noise and the effectiveness of corrective measures were clearly defined by laboratory tests. These findings were confirmed by the third series of tests, with noise measurements recorded on the rectifier locomotives and at Quarryville. A general summary of significant road tests with filtering arrangements and maximum noise readings observed at Quarryville is presented in Table I.

During the course of these tests, observations were made of noise to ground which resulted as other types of electric locomotives operated through the same exposure. Pertinent data are shown in Table II. These values do not necessarily represent a true comparison of the various types, as complete information on their operation was not available.

Some of the data obtained during the first series of tests have been plotted with respect to time in Fig. 4, which shows movement of trains between Safe Harbor and Parkesburg, current drawn from the overhead wire by the 2-unit rectifier locomotive, noise measurements taken on the locomotive, and noise measurements obtained on the railroad telephone circuit at Parkesburg. Examination of the last-named values with respect to locomotive current indicates clearly the influences from other electric locomotives, and shows that the capacitors on the rectifier locomotive are effective in filtering out other noises as well as noise originated in the rectifier.



1 CAPACITOR/MOTOR 3 CAPACITORS/MOTOR 5 CAPACITORS/MOTOR

Fig. 6. Data obtained from third series of locomotive tests, showing corrective effects of capacitors

Significant data from the second series of tests are plotted in Fig. 5. In so far as noise measurements obtained at Quarryville are concerned, influence can be attributed to operation of railroad motive power only when trains are moving between Providence and Bart. Although it is evident that noise directly chargeable to the rectifier locomotive without filters is considerably higher than that resulting from an older type of electric locomotive (class *P5a* locomotive 4703), it is also obvious that noise is experienced in this telephone circuit from other sources. The sound levels indicated on these graphs represent the readings of a sound-level meter measuring the noises emanating from the tape recorder loud speaker. All readings were obtained by placing the meter on the same table top 12 inches in front of the loud speaker.

The corrective effects of various numbers of capacitors connected across each traction motor are illustrated in Fig. 6. Noise values somewhat higher than those measured with the older electric locomotives are shown with three capacitors per motor; values are somewhat lowered with five capacitors per motor. Analysis of the *I·T* products, as measured on the locomotives, indicates that relatively little improvement is realized as the number of capacitors is increased above three. Adjustment of damping resistor values would make the 3-capacitor filter more effective.

Discussion of Test Data

During the road tests, there were no means of blanking out the noise effects being produced in the telephone circuits by other electric locomotives which were operating simultaneously from the same overhead system, or from noise influence of adjacent power circuits, signal circuits, etc.

It was evident that the single-phase series motor generated many more frequencies than were generated by the single-phase rectifiers. It was also known that the test telephone circuits were exposed to 60-cycle power circuits and 100-cycle signal circuits, as well as to the 25-cycle electrification system. Analysis of the sound records showed the presence of 60 cycles and some of its harmonics as well as 25 cycles and its harmonics in the telephone circuit.

It was necessary to search for means during the studies and analyses to distinguish between the noise effects of the rectifier locomotive and the noise effects of series motors and other noise ambients in the test telephone circuits. Some of these noise effects were rather difficult to differentiate from the characteristic noise from the rectifier locomotive, as described later. This was particularly true of the 60-cycle hum and the noise that resulted when the single-phase series motor was worked above the knee of its saturation curve.

Four general measuring methods have been used on the sound records, as follows:

1. Sound effects by "hearing" the noises emanating from the loud speaker.
2. "Sight" effects by cathode-ray oscilloscope pattern of voltage across the loud-speaker coil.
3. General Radio sound level meter readings of the sound from the loud speaker.
4. Harmonic analyses of voltages across the loud-speaker coil by a General Radio harmonic analyzer.

Data obtained in the early tests were not suitable for quantitative analysis, but were suitable for qualitative analysis.

The following filters for limiting the voice frequency band input to the loud speaker were provided: 1. 300 cycles per second (cps) high-pass filter; 2. 600 cps high-pass filter; 3. 3,500 cps low-pass filter. The recorder had a high-frequency tone control dial that was used in a predetermined manner during the studies. Use of these filters permits determining whether most of the noise volume is being produced in the lower or higher frequency ranges of the voice frequency band.

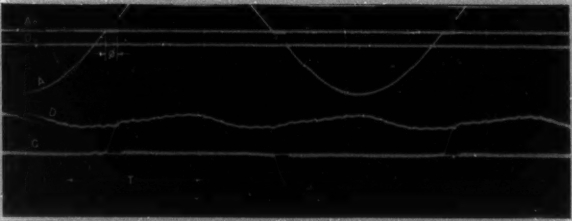


Fig. 7. Oscillogram showing motor performance with 50-per-cent field current and 400 amperes in trolley

A—Anode-to-anode voltage. Peak is 1,090 volts
B—D-c motor current. Peak is 360 amperes. Ripple is 8.45 per cent
C—Air core coil volts in 783 turns. Peak is 11 volts
ϕ—Angle of overlap, 13.2 degrees
T—One cycle with 60-cycle timing

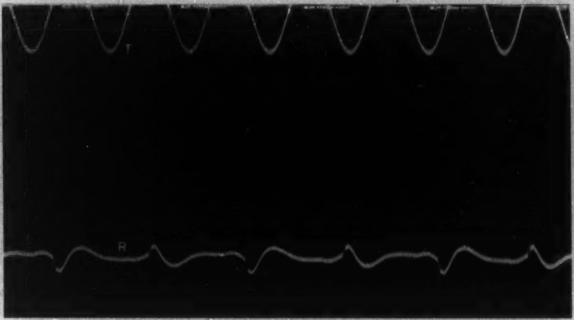


Fig. 8. Oscillogram showing recorder output volts with 500 amperes in trolley and without telephone filter

T—60-cycle timing wave

R—Recorder output volts

Noise—Rectifier Locomotive

The most predominant and most characteristic sound produced by a nonfiltered single-phase rectifier locomotive in exposed communication circuits is a "buzz," similar to the sound of a small buzzer. The electromagnetic characteristics of the longitudinal coupling between the single-phase catenary circuit and an exposed communication circuit are similar to those of an air core current transformer. The buzz in the communication circuit is induced by the high rate of change in the catenary current and flux during the angle of overlap. When tube no. 2 (Fig. 3) fires with tube no. 1 still conducting, an incipient short circuit is set up through the transformer windings because of energy stored in the reactance of the circuit. This angle of overlap starts when tube no. 2 starts conducting and ends when current stops flowing in tube no. 1. While both tubes are conducting, the increasing current from tube no. 2 flows through tube no. 1 in the opposite direction to the latter's current, which is decreasing. At the instant these two opposing currents become equal, tube no. 1 stops conducting. When the current in tube no. 1 ceases, there is still some value of flux, air core, and iron present caused by the tube no. 1 current, which must collapse at a high rate of speed.

The longitudinal voltage induced in

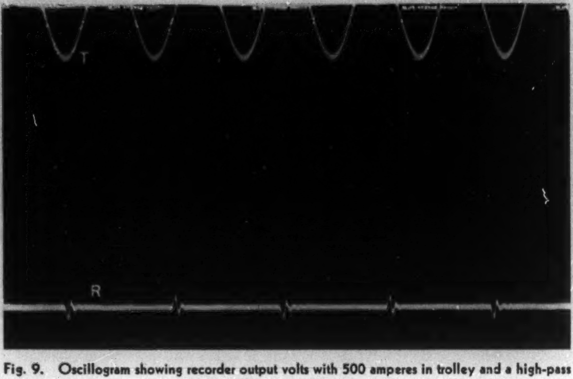


Fig. 9. Oscillogram showing recorder output volts with 500 amperes in trolley and a high-pass telephone filter with 300-cycle cutoff

T—60-cycle timing wave R—Recorder output volts

the exposed communication circuit shows up as a pulse in voltage during the angle of overlap. The value of induced longitudinal voltage varies directly with the load current being commutated and inversely with the angle of overlap. The pulse in longitudinal voltage in the communication circuit will occur every half cycle, as shown in Fig. 7. It will be positive and negative during alternate half cycles.

Since the communication circuit is not perfectly balanced, some percentage of this induced pulse of longitudinal voltage will appear across the connected receivers, and this reduced pulse of voltage will cause the receiver diaphragm to pulse each half cycle.

Rectifier Buzz Setup on Test Floor

A laboratory test setup was made to approximate the buzz from the rectifier locomotive line current at Quarryville to determine how the buzz varied with angle of overlap, load current, filter capacity and other damping means, and with $I \cdot T$ as measured by the noise meter. Coupling was made by means of an air-core current transformer. Current in the through lead was varied from 100 to 500 amperes to duplicate the line current drawn by one rectifier locomotive unit.

This line current was measured on a low-voltage primary winding so the exciting current component of the line current was much greater than that on the 11-kv winding of the rectifier locomotive transformer. The exciting current was approximately 111 amperes in the test setup, whereas the exciting current of the 11-kv winding on the locomotive was approximately 6 amperes.

The voltage and reactance of the test circuit versus the capacitance of one to ten capacitors were proportioned, so that one to ten capacitors had approximately the same relations as on the locomotive. Hence, any number of capacitors in the test floor setup should approximate the same filtering action as the same number on the locomotive.

The angle of overlap for a given value of line current on the test floor had to be made somewhat greater than the angle of overlap for a corresponding value of line current on the locomotive, and the capacitors had to be applied at a reduced kva rating in order to obtain these proportions. The frequencies that would be set up in a telephone circuit by such an exposure were registered on a tape recorder.

Fig. 8 shows the frequencies that are transmitted through the air core coupling to the tape. Some of the low-frequency disturbance can be charged to the odd harmonics in the exciting current. The "spike" caused by the rectifier produces the disturbance that results in the buzz in a telephone receiver. When this spike is eliminated, the buzz disappears. Fig. 9 shows the same record when it is filtered with a 300-cps high-pass filter. Most of the low frequencies have disappeared but the spike is still there.

The old point of view, based on the square-wave analytical approach to the single-phase rectifier, involved a large mass of odd harmonics which was visualized as flowing continuously in the catenary circuit and exposed telephone circuit, as when generated by series motors, etc. However, Fig. 7 shows none of these frequencies, and tests indicate that only the spike need be eliminated to remove the buzz.

This spike occurs at the end of the angle of overlap, and it can be seen in Fig. 9 that it covers only a small percentage of each 180 degrees. Hence, the spikes from two adjacent units can be kept from adding to each other if the line currents are dephased a few degrees. Spikes from motive power units on different trains are inherently dephased by constantly varying reactance between the rectifiers and the generators which supply electrification systems.

Conclusions

Rectifier motive power does not involve a telephone noise problem which can not be successfully and satisfactorily co-ordinated. It has been demonstrated by tests in the laboratory and on the railroad that a nonfiltered, high-capacity, single-phase rectifier locomotive will produce noise in exposed communication circuits, but that this noise can be diminished to satisfactory values by comparatively small and simple filters on the locomotive.

◆

Discussion

H. F. Brown (New Haven, Conn.): It is apparent that there may be a considerable number of motive power units of the type described applied in the near future to existing 25-cycle electrified lines. Can the authors offer any opinion as to whether the disturbing effects of a number of locomotives or cars of this type operating at random points on a system will tend to be additive or not? If so, would resonant shunts, tuned to the range of audio frequencies set up, distributed along the line particularly at locations of severe exposure, be as effective a solution as such filtering equipment carried on the locomotive?

Considering the possibility that the rectifier locomotive may be applied to commercial frequency (60-cycle) railway electrification in this country in the future, would not this change to a higher fundamental frequency tend to reduce the inductive co-ordination problem, since the current in the trolley contact line could be reduced to at least one-half of the values which now exist on the 25-cycle systems?

So far, these tests have included only the inductive coupling of the 25-cycle railway contact system and the paralleling communication lines. Would the mitigation problem be any more difficult of solution where the harmonics could be put back into the general power network supplying not only commercial frequency to the railway motive power but to all other communication apparatus such as radio and television?

Does the inherent impedance of the power system supplying the power to the trolley contact system have any appreciable effect on the *I·T* factors? That is to say, would the *I·T* factors for a given locomotive be the same operating on a large system having a low impedance between the generation and the locomotive as on a smaller system having an impedance of at least twice that of the larger system? The concluding paragraph of the paper should be very encouraging to potential future users of this type of motive power.

F. T. Garry (Southern New England Telephone and Telegraph Company, New Haven, Conn.): It has been pointed out that ignitron-type multiple-unit cars and locomotives develop *I·T* products of sufficient magnitude to cause objectionable levels of noise in paralleling communication circuits, if such units are not equipped with filters of proper design.

In the exposure areas that were chosen for test purposes, the data recorded indicate that the noise to ground in the disturbed circuits is in direct relationship to the *I·T* product of the ignitron unit in operation. The effect of various combinations of capacitors in reducing the *I·T* product developed has been shown together with the effect produced by the introduction of a dephasing reactor. These tests indicate that with a filter the *I·T* product, and hence the noise to ground in the telephone circuit, can be materially reduced. The conclusion seems logical that the introduction of this type of locomotive or MU car should not present an inductive co-ordination problem that cannot be satisfactorily resolved.

When the New York, New Haven and Hartford Railroad ordered ten ignitron locomotives and 100 MU cars for use, primarily on their main line between New York and New Haven, both their communication engineers and those of The Southern New England Telephone Company were concerned with the possible increased levels of noise on their circuits. While it was known that filters could be applied to reduce the *I·T* products of such equipment, two complications arose in the subject of filter application. The first, applying to the ignitron locomotive, is the weight limitation imposed by the entrance to Grand Central Terminal in New York City. Here we have a possible limitation to the amount of filtering strictly from the viewpoint of the additional weight imposed. A question also arises as to the cumulative effect of a combination of locomotives and MU cars operating in the same vicinity. The New Haven Road will be using many of the MU cars in tandem. How the influence of the various units will add in effect on communication circuits is probably a matter that will have to be resolved by tests. The present operating locomotives have not been a noise problem to the communication facilities involved.

In the New Haven Railroad area, we are not certain that the noise in the paralleling communication circuits will be at a satisfactory level after the tentatively scheduled filtering mentioned in the report. For example, (granting that the exposures are not the same) although some of the metallic noise levels shown in the report would be satisfactory at the switchboard end of a trunk circuit, they would be objectionable at that level in the subscriber's receiver. We know from experience that metallic noise in a circuit that results in a receiver noise greater than 17 to 20 db may result in subscriber reaction.

Mr. Brown indicates that consideration is being given to the use of ignitron-type locomotives or MU cars on railroad lines where 60-cycle a-c voltage is supplied from commercial power systems. Can filtering be applied in such a manner as to confine the harmonic frequencies generated to the railroad area or will they be present, to a marked degree, in the 60-cycle source? This problem is one of vital concern to telephone companies because of present exposures to such commercial power circuits and to a constant increase in the joint use of plant with higher voltage distribution circuits.

D. R. MacLeod (General Electric Company, Erie, Pa.): One of the major items to be considered in estimating the cost of an electrification is the cost of inductive co-ordination with the railroad's own communication circuits. The general policy of the commercial communication companies of considering inductive co-ordination with potential railroad electrifications in all their planning, and the joint agreement that they made with the Association of American Railroads nearly 25 years ago, have been major contributions to the cause of a-c electrification in this country. This paper is further evidence of the excellent spirit of co-operation that exists between the railroads, the communication companies, and the electrical manufacturers.

The paper gives certain qualitative data in regard to the influence of single phase rectifier-type motive power and makes comparisons with certain other types of single-phase motive power on the Pennsylvania Railroad electrification. It should be borne in mind that the wave shape variations of a rectifier type of equipment are different from those of an a-c series motor type. The telephone influence of a-c series motor type of locomotives such as the *P-5A* and *GG-1* is a result of the commutator ripple and slot harmonics. The telephone influence factor (tif) varies with the speed of the motors. On the other hand, the tif of a rectifier locomotive depends on such things as the amount of smoothing in the d-c traction motor circuits, the current drawn by the motors, etc., so that its tif does not vary with speed. The rectifier-type MU car has wave shape characteristics similar to those of the locomotive. The motors on car 764 mentioned in the paper are the obsolete doubly fed type of motor which has very poor wave shape characteristics as compared with modern a-c traction motors used on equipments purchased by the railroad during the last 25 years. It is my opinion that the data given in the paper for a-c multiple-unit cars should not be taken as representative of modern a-c MU car equipments. There is a large number of cars with doubly fed motors in operation on the Pennsylvania Railroad, and the fact that cars with an *I·T* of 2,000 have operated in 10-car trains long before modern methods of shielding communication circuits were available is of significance in considering the application of all types of rectifier motive power.

When applying rectifier motive power to a system where only the a-c series motor type of motive power has been used before, it is necessary to know the characteristics of the latter in some detail. It is hoped that the following data will add to the value of the paper.

In January 1931, a series of tests was made at the Erie Works of the General Electric Company on a sample motor that was used for the *P-5A* locomotives. This type of motor has been described.[1] Tests were made to determine whether it was necessary to use skewed slots on a-c traction motors. The conclusion was that there would be very little difference in the telephone influence of a locomotive having motors with skewed slots and one having motors with straight slots. The motors were built with straight slots and the data given below refer to this type of armature. The tif meter used in the tests was weighted in accordance with the 1919 weighting curve.[2] The values obtained are therefore higher than if a meter with the 1935 weighting had been used.

The tests were made by holding the current constant as the speed was varied from 80 to 1,000 rpm. The maximum speed of the motor is 1,250 rpm. Currents of 1,000, 1,500, 2,000, and 2,500 amperes were used. The maximum sustained current used on *P-5A* locomotives in full field is 4,000 amperes. The variations in tif readings at a given motor speed for the four currents tested were so small that it is a good approximation to assume that the *I·T* (current times tif) of an a-c series motor locomotive for a given locomotive speed varies directly with the input current. Results were obtained for 1,000 amperes through the motor during the 1931 factory tests, as shown in Table III.

P-5A locomotives haul tonnage trains over the grade in the vicinity of Quarryville at speeds between 30 and 40 miles per hour. It is only if a signal is against them or if the rail conditions are bad that sustained speeds as low as 15 miles per hour (mph) are held by these locomotives anywhere on the electrification. A *P-5A* locomotive takes approximately 460 amperes from the 11-kv trolley when hauling rated tonnage trains up the 0.3 per cent grade. It takes 230 amperes when hauling these trains at 15 mph. Since the tif increases more than the current decreases, it is possible that higher *I·T* products than those recorded in the

Table III. Results of 1931 Factory Tests

Motor Rpm	*P-5A* Locomotive Speed, 70 Mph Maximum	Tif 1919 Weighting
80	4.4	12
100	5.5	16
115	6.4	18.5
125	6.9	10
150	8.3	6
200	11.0	10
225	12.4	19
250	13.8	11
300	16.5	5.5
500	27.6	2
700	38.6	1.5
900	49.6	4.0
1,000	55.2	2.5

Table IV. Frequencies Identified and Measured

Frequency	Amperes	% of Fundamental
25	1,500	100.0
175	1.042	0.070
225	0.704	0.047
311	0.315	0.021
375	0.27	0.018
575	0.318	0.021
625	0.267	0.018
725	0.182	0.012
752	0.174	0.012
887	0.145	0.0097
1,100	1.18	0.079
1,150	1.79	0.120
1,675	0.106	0.0077
2,250	0.98	0.065
2,300	0.819	0.055
3,380	0.526	0.035

paper would have been obtained for both the *GG-1* and *P-5A* locomotives if they had been tested at slow speed.

The tests that were made with the wave analyzer in 1931 showed that at 425 rpm, 84 per cent of the total tif was contributed by the 1,058-, 1,108-, 2,145- and 2,195-cycle components. With the motors operating at about 220 rpm, 99 per cent of the tif was contributed by the 1,100- and 1,150-cycle components. Frequencies were identified and measured by the wave analyzer with the motor operating at 220 rpm with 1,500 amperes through it, as shown in Table IV. The calculated tif with the 1919 curve works out at 23.1, whereas the measured tif using the 1919 weighted meter was 19.0.

The conclusion reached in the paper that rectifier motive power does not involve a telephone noise problem which cannot be successfully and satisfactorily co-ordinated is a very valuable contribution. I should like to ask the authors what $I \cdot T$ value they consider necessary for a fleet of locomotives of this type to operate anywhere on the Pennsylvania Railroad electrification without additional changes in the wayside installations.

REFERENCES

1. THE MODERN SINGLE-PHASE MOTOR FOR RAILROAD ELECTRIFICATION, F. H. Pritchard, Felix Konn. *AIEE Transactions*, vol. 50, March 1931, pp. 263–68.

2. REVIEW OF WORK OF SUB-COMMITTEE ON WAVE SHAPE STANDARD OF THE STANDARDS COMMITTEE, Harold S. Osborne. *AIEE Transactions*, vol. 38, pt. I, 1914, p. 265.

C. W. Frick (General Electric Company, Schenectady, N. Y.): The authors are to be congratulated for their interesting paper on a timely subject. As they point out, one phase of this inductive co-ordination study is the evaluation of the inductive influence of the rectifier motive power equipment used on locomotives and cars. I should like to comment on some aspects of this subject, particularly the buzz type of noise.

The test conditions described in the paper include commercial telephone systems and railway telephone systems. The latter systems often use amplifiers and loud speakers. The paper mentions that the buzz type of noise was first observed on railway circuits which may have been so equipped. Recurrent voltage spikes impressed on electronic equipment such as amplifiers and radio receivers produce a buzzing sound as described. However, we do not believe the effect on commercial telephone equipment would be very pronounced since the spikes would tend to be smoothed out in the process of transmission from the source to the telephones that might be affected.

The spikes of voltage which are induced by the line current in the secondary of an air transformer are accounted for easily by analysis. The line current may be treated as made up of two parts. During the commutating period the current is made up of the short-circuit current which would be produced by short-circuiting the rectifier transformer, combined with a d-c component. After commutation is completed, the current is substantially a steady direct current until the next commutation starts. The only component which induces voltage is the short-circuit component which exists only during the commutating period. The short-circuit current is a cosine wave of the fundamental frequency of the power supply. Knowing the current and the mutual inductance the induced voltage can be calculated. It is easily shown that the induced voltage starts at zero, rises to a maximum, and then drops to zero at the end of the commutating period where it remains until the next commutation. This is the curve of the spike voltage shown on Fig. 7. The effect is produced by rectifier operation with any number of phases at any frequency and is not peculiar to single-phase operation at 25 cycles. The only difference is the slower rate at which the spikes occur, 50 per second, as compared with 360 per second for 6-phase operation at 60 cycles.

$I \cdot T$ has been found to be a reliable index to the influence on commercial telephone systems of rectifier current for the usual power rectifiers. In the tests reported in the paper, the noise-to-ground produced by operation of the rectifier locomotive and measured with the noise set, shows good correlation with the $I \cdot T$ values. It is easier to make the comparison when the noise values and the $I \cdot T$ values are in the same units, such as db. One of the authors of the paper has explained that the curves on Fig. 4 and Fig. 6 marked "2-unit db" are actually curves of $I \cdot T$ in db. These curves are easy to compare with the noise curves which are also in db. It was also mentioned that 51 on this db-scale corresponds to an $I \cdot T$ of 24,000. Using this reference point, some of the $I \cdot T$ values tabulated in the paper have been converted to decibels as shown in the tabulation which follows. The $I \cdot T$ values were opposite the noise-to-ground values except those in parentheses which were found in another part of the paper and are believed to be applicable. Refer to Table V.

Table V. $I \cdot T$ and Noise-to-Ground Values

Rectifier Locomotive	$I \cdot T$	$I \cdot T$, Db	Noise to Ground, Db
Unfiltered	24,000	51	51
Reactor	12,000	45	44
10 capacitors	3,400	34	39
10 capacitors and reactor	2,200	30	35
Two *P5A* locomotives	(2,800)	32	36
Two *GG1* locomotives	(4,500)	36	33

Since the $I \cdot T$ shows reasonable correlation with measured noise and the noise set was designed to represent the characteristics of commercial telephone equipment, it seems reasonable to expect that the $I \cdot T$ of rectifier motive power current is a reliable index to telephone influence, at least for most commercial telephone systems. This of course may not be true for other wire communication systems such as railway systems with amplifiers and loud speakers.

As to determining noise by the combination of harmonics, I believe that the method would work even in the case of noises of the buzz type provided the method is properly applied. Very likely the effect of the noise pulses (spikes) on an amplifier and loud speaker involves shock excitation, which would appear to complicate the analysis. However, the same situation is encountered when we analyze the effect of the steep wave fronts of square waves or recurrent pulses of very short duration on a radio receiver. The buzz is evident at pulse repetition rates up to several hundred per second. I have made analyses of the effect of steep wave fronts and pulses on a receiver, or radio noise meter, starting with the square wave or pulse represented by a series of harmonics. The results check quite well with the noise meter readings.

If a new method of measurement is needed to evaluate, for example, the influence of rectifier motive power on loud speaker telephone systems, a different metering system from the one used on present noise measuring sets might be needed. I do not think the sound level meter to measure the output of the loud speaker would solve the problem. The sound level meter responds to average power, while in this case the ear responds to the short pulses. To develop a metering system on a firm basis it would seem necessary to start with an evaluation of the effects on groups of observers. This is one phase of the procedure by which the noise measuring set for commercial telephone systems was developed. In view of the rather limited application, such an elaborate procedure as this does not seem worth while, especially when the present situation may be taken care of by the application of reasonable filtering to the rectifier motive power.

H. S. Ogden (General Electric Company, Erie, Pa.): In the paper a statement is made that a harmonic wave analyzer was used in the tests but no information is given on the results of the analyzer measurements. It would have been most helpful had the wave analyzer results been included since a wave analysis with the percentages of the various harmonics is of great value when making comparisons with telephone interference weighting factors.

In dealing with the dephasing reactor this paper indicates that the later tests were made without the benefit of this device. Table I shows but a small and inconsistent benefit to be derived from the device if one considers the column headed "Metallic Noise." Presumably this is one of the reasons that tests with the reactor were not repeated.

Under "Noise-Rectifier Locomotive," an incipient short circuit is referred to that is set up when a rectifier fires. There is nothing incipient about it; it is very real, and there is a continuing short circuit known as arc back if a tube does not seal off when the current falls to zero and it tries to conduct in the opposite direction.

The paper lacks oscillograph pictures of the primary ampere and voltage waves to the locomotive. Since such elaborate setups were made to obtain this basic information, it might well have been included.

Some United States patents assigned to the Westinghouse Electric Corporation describe a tuned second harmonic filter circuit for use in the traction motor and main smoothing reactor circuits. Such a second harmonic filter circuit should have an effect on the harmonic content of primary current wave. Thus it should have some effect on the spike that the authors describe. Nowhere in the paper is any reference made to this system. Does such a system have any beneficial effect on telephone interference problems?

F. D. Heiss (Bell Telephone Company of Pennsylvania, Philadelphia, Pa.): The tests of the noise influence of the rectifier locomotive, constitute an excellent example of the co-operative process essential to the solution of many problems of noise or low-frequency induction. Teamwork among the several interested parties was practically perfect. The word interested is used advisedly. The telephone group was, of course, somewhat fearful as to the possible future inductive effects on a rather wide geographical basis of this new unknown. The railroad company was interested, both as owner of a possible noise source which might cause the telephone companies to come to it with grievances, and as operator of its own extensive communication network which might be even more seriously affected than the telephone companies' circuits. Finally, the manufacturer was interested in the production of a device that would not produce future repercussions from purchasers or communication companies.

Much careful planning was required before the tests could be made. Availability of the locomotive for installation of test equipment was a problem, but this was only the first step. Many telephone calls were needed to determine when the locomotive had returned to the location (Harrisburg) established for start of the test runs; whether and when it would probably be next used; when it was finally and definitely scheduled to leave; when it had actually left; and its progress over the some 150 miles of route. Then more calls were needed to get the test personnel to the locomotive and to the field test points at the proper time. Freight train schedules are intentionally made much more flexible than passenger train schedules. Locomotives are placed in a pool rather than pre-assigned some time in advance. This very mundane part of the testing co-ordination required many man-hours of effort.

Selection of locations for noise measurements on railroad and telephone company circuits requires some care. Not only is the type of plant (open-wire, aerial or underground cable) and the extent and nature of the exposure, including shielding by other plant or structures, of prime importance, but consideration must also be given to normal train operation at the location. If the exposure is along a railroad grade, or at a point where trains are normally starting with heavy loads, the current drain and resultant noise influence may be much more severe than at other locations. At one of the test points chosen, it developed that on only one run was sufficient trolley current drawn to produce satisfactory test data.

As a result of the co-ordinated tests, that with only two locomotives (four units) operating, it appears that a simple filter consisting of only three capacitors per motor will satisfactorily reduce the noise influence. However, if a large portion of the total motive power were provided by ignitron locomotives it would probably be desirable to ensure that the resultant influence would not exceed that caused by the a-c locomotives of the several types that have been in use for some time. In this connection, the experience with a-c locomotives over a number of years has been that no serious problems have arisen from a noise standpoint.

C. C. Whittaker (Westinghouse Electric Corporation, East Pittsburgh, Pa.): This paper is of interest to proponents of electrification in that it offers a solution of the last remaining problem having to do with locomotive design as influencing the economics of electrification of railroads.

In our 50 years of electrification experience we have learned that the best way to transmit power to the trolley wire is by high-voltage alternating current, meaning 11,000 volts, 24,000 volts, or even 48,000 volts.

Because of the wide interest in the Diesel electric locomotive, we now appreciate more fully the advantages of having all locomotive weight on drivers to obtain maximum tractive effort. We have learned also that small wheels, in simple truck designs, result in ease of maintenance to both traction motors and trucks.

The inherently simple d-c series traction motor, highly tooled for assembly line production, resulting in a low-cost motor, has also sold itself. The d-c motor has likewise found favor with the railroads because of its relatively flat speed-tractive effort characteristic. Because of this, should slipping occur, a d-c series motor loses tractive effort at a high rate for a slight increase in motor speed, compared with the similar a-c motor speed-tractive effort characteristic. This feature, with parallel-connected motors, has proved its worth in minimizing initial slipping and avoiding spinning after slipping occurs. If motors are connected in series, this advantage disappears.

During the past $4^1/_2$ years we have proved by actual railroad service that the ignitron rectifier is a very satisfactory means of conversion of power from the a-c line to the d-c motor. The one question which for a time was unanswered, concerned the influence of rectified power on communication lines. We feel that this last question has now been answered by this paper. Test results presented show that interference can be controlled with an acceptable amount of filter equipment. We thus draw the conclusion that new electrifications, preferably at 60 cycles, may now be undertaken, with confidence that motive power units in a broad sense have reached a mature development, and that long life and reasonable maintenance are assured.

E. B. King, K. H. Gordon, L. J. Hibbard: The authors wish to acknowledge with thanks the many helpful discussions on this paper. In reply to Mr. Brown and to Mr. Garry, the inductive influence effect of a number of locomotives and cars operating at random points on the system can never be additive. Differences in reactance and loads with intervening substations, etc., will keep the various buzz spikes dephased with respect to each other.

Increased voltage on the overhead catenary conductor along with a corresponding decrease in the current for a given frequency and a given kva load will reduce the trolley $I{\cdot}T$ product. The use of 60 cycles in place of 25 cycles for the same set of conditions will increase the trolley $I{\cdot}T$ product because the telephone receiver and the human ear are much more sensitive to the lower harmonics of 60-cycle current than to those of 25-cycle current.

There is some higher value of 60-cycle trolley voltage with the corresponding reduced value of trolley current that will give the same $I{\cdot}T$ product on 60 cycles as is obtained on 25 cycles with existing operating voltage and current conditions.

The mitigative problem would definitely be more difficult if the wave shape of the 60-cycle supply were affected by direct connection of commercial power sources to railroad rectifier motive equipment. However, when these new problems arise, we feel confident that a satisfactory solution to the inductive co-ordination aspects will be determined by co-operative studies as have been made in the past.

Increased reactance from anode to anode in the a-c circuit increases the angle of overlap, which in turn reduces the rectifier buzz. The anode to anode reactance is increased as the system reactance increases. However, most of the anode-to-anode reactance is usually governed by the locomotive transformer, anode reactors, etc., and the system reactance reflected to the anode-to-anode voltage base is only a small per cent of the total anode-to-anode reactance.

Replying to Mr. MacLeod, the statement is made that a tif meter weighted in accordance with the 1919 weighting curve would result in higher influence values than for a meter using the 1935 weighting. The authors agree that this would be true if the prominent harmonics should fall in the range between 1,000 and 1,500 cycles, or below 750 cycles. However, if the prominent harmonics were outside this range, the reverse would be true. Where many frequencies are involved, it is necessary to take the root-sum-square of all the individual weighted harmonics.

As to the question regarding the $I{\cdot}T$ value considered necessary for a fleet of locomotives of this type operating anywhere on the Pennsylvania Railroad electrification, it is impracticable to specify such a limit at this time. Further tests would be necessary to determine the additive effects of many rectifier units distributed over the system.

In reply to Mr. Frick, and referring to Table V, which correlates the $I{\cdot}T$ value with the noise-to-ground, the use of the assumed values of $I{\cdot}T$ in parentheses for two a-c types of locomotives is questioned. Apparently the values shown in the paper for single unit locomotives have been doubled. However, the single locomotive values were measured at other locations under different conditions and it is likely that the $I{\cdot}T$ for the 2-unit locomotives would be much lower than the indicated value. We believe that the correlation is well demonstrated without the use of these two examples.

In reply to Mr. Ogden, regarding the omission of analyzer measurements, it was impossible to obtain such an analysis on a moving locomotive under power because the load and circuit conditions were continually changing. For such an analysis to have any significant meaning, it would require that the base load and circuit conditions be held absolutely constant while all harmonic readings are being taken.

Dephasing reactors in addition to filters would be considered in cases of extreme exposure conditions, since the dephasing of the buzz spikes reduces the magnitude of the noise problem.

The second harmonic filter is a d-c filter which permits a higher ripple to be used in the d-c circuit. Increased ripple in the d-c circuit tends to reduce the buzz spike and the over-all inductive influence. In fact, if a single-phase rectifier feeds a pure resistance load with zero reactance in the a-c circuit and zero inductance in the d-c circuit, the buzz spike would not occur.

An Investigation of the Switching Criteria for Higher Order Contactor Servomechanisms

IRVING BOGNER
ASSOCIATE MEMBER AIEE

LOUIS F. KAZDA
MEMBER AIEE

Synopsis: The use of a linear anticipatory switching network in an on-off servomechanism, while improving system stability, does not result in optimum response. Mathematical analysis of second order systems indicates that the network should be a nonlinear function of error and error rate. This paper describes a method for deriving the switching criterion for a "piecewise" linear on-off servo of order greater than two, when the input variable is limited to a step function of position and velocity. The switching criterion results in system error and in derivatives of error reducing to zero in a minimum time. System behavior is described in the principal co-ordinate phase space. Starting at a point in the phase space corresponding to the initial conditions, the representative point moves in a series of discontinuous trajectories determined by the sign of the controlled variable, eventually reaching the origin. The sequence of sign reversals represents the desired switching criterion.

Results indicate that: 1. a unique switching criterion is possible when the number of sign reversals is one less than the order of the system; and 2. although any arbitrary number of reversals greater than the above minimum may be used, in the examples considered minimum switching resulted in minimum response time. All the recorded data and photographs were obtained from an analogue computer study.

FEEDBACK control theory may be divided into two major categories, depending upon the type of control that exists. The continuous control type of servo is defined by Brown and Campbell[1] as being "linearly and continuously error sensitive." Such systems, described by a linear differential equation with real constant coefficients, are discussed extensively in the literature. The assumption of linearity permits one to employ methods of analysis not applicable to nonlinear systems. Since the linear servo is in many respects similar to a feedback amplifier, the frequency response methods of Nyquist[2] and Bode[3] represent powerful tools in linear design.

The characterization of a physical system as linear has resulted in the development of fast and accurate servos, which are satisfactory in many applications. Frequently, however, control systems possess nonlinearities, inherent or inserted, which are not linear in the sense that they can be described by a linear differential equation.

The past decade has brought forth important contributions to the study of nonlinear servomechanisms, particularly in the field of contactor or on-off servos. In the earlier analyses by MacColl[4] and Weiss,[5] phase plane[6,7] techniques are used to determine the responses of second-order on-off servos possessing such nonlinearities as finite torque reversal time and coulomb damping. Kochenburger[8] and Johnson[9] modified existing linear frequency response techniques to analyze nonlinear systems.

Emphasis in much of the literature pertaining to contactor servos has been on methods of improving system stability. Exceptions are the methods suggested by Neiswander and MacNeal,[10] and before them by Hopkin[11] and McDonald,[12] in which phase plane analysis is used to optimize performance. In McDonald's method a torque-reversing criterion is developed in which the sign of the torque applied to the load is determined by a nonlinear function of error and error rate.

It is the purpose of this paper, in extending the suggestions put forth by McDonald, to present a method for obtaining unique switching criteria for contactor servomechanisms of order greater than two. The analysis presupposes a discontinuous forcing function, fixed in magnitude and reversible in sign, supplied to a dynamic system which is describable over every interval, of its operation by a linear differential equation. This type of control, while nonlinear in the general sense of the term, is more accurately described as being sectionally or piecewise linear. The discussion which follows makes use of phase plane and phase space analysis in describing the system behavior, the particular phase space of interest being one in which servo error is plotted against its higher derivatives.

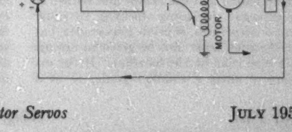


Fig. 1 (right). Second-order contactor servomechanism. The voltage v reverses in sign, as a nonlinear function of E and (dE)/(dt) in a manner which reduces E_0 or $[(dE)/(dt)]_0$ to zero in a minimum time. The motor field time constant is neglected

---

Paper 54-44, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing November 17, 1953.

IRVING BOGNER is with the Cook Research Laboratories, Skokie, Ill.; LOUIS F. KAZDA is with the University of Michigan, Ann Arbor, Mich.

The authors wish to acknowledge the assistance received from the engineering faculty at the University of Michigan and the staff at Cook Research Laboratories.

This work in part was sponsored by the United States Air Force under Contract No. *AF 33(038)-21673*, and contains part of the results of a thesis submitted by Irving Bogner in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Michigan.

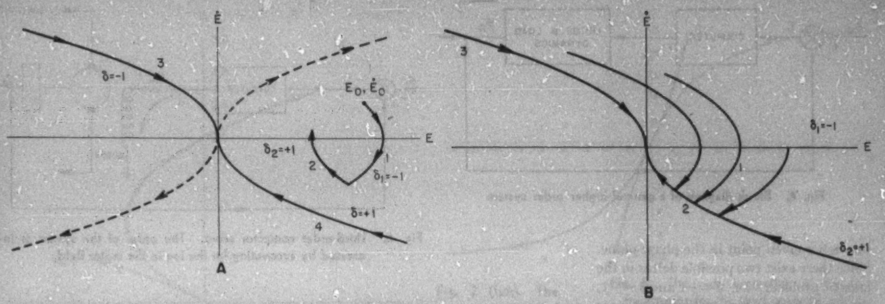


Fig. 2. Phase portraits for the second-order contactor servo of Fig. 1. Curves 1 and 3 in Fig. 2(A) represent trajectories for a positive value of v; curves 2 and 4 for a negative value. Curves 2 and 3 in Fig. 2(B) are the sign reversal curves for optimum system response to step inputs of position and velocity.

Subsequent analysis makes use of what is referred to in the literature as the principal co-ordinate[13] phase space.

Definition of Symbols

a = number of controlled variable sign reversals in exeess of the minimum required
b_n = real coefficients in the differential equation describing system dynamics
C = constant of integration
E = servomechanism position error
E_0 = servomechanism position error at time equal to zero
ϵ = dimensionless position error
$F(\quad)$ = function of $(\quad)$
$h_n(\quad)$ = equation of a trajectory projection
i = motor field current
I_a = motor armature current, assumed constant
J = moment of inertia of motor and load
K_ϕ = ratio of flux to field current
K_2 = ratio of torque to flux
L = inductance of motor field
P = number of sign reversals of controlled variable
R = resistance of motor field
S_n = switching curves defined by equations 45 and 46
t = time
T = motor torque
T_M = motor saturation torque
u_n = phase space co-ordinates corresponding to $\frac{d^{n-1}E}{dt^{n-1}}$
v_n = principal co-ordinate of a system
v_{12} = variable v_1 at point 2 in the phase space
v = voltage output of the servomechanism computer, the controlled variable
V = magnitude of v
δ = sign determining factor for the controlled variable
Δ = increment of a variable
θ_i = servomechanism input variable
Θ_i = dimensionless input variable
θ_0 = servomechanism output variable
Θ_0 = dimensionless output variable
$\sum_n$ = summation on n
τ = dimensionless time
τ_ϕ = motor field time constant
ϕ = field flux

Optimum Switching for Second-Order Systems

For the idealized system shown in Fig. 1, the load differential equation of motion is

$$\pm T = J\frac{d^2\theta_0}{dt^2} \tag{1}$$

Under the assumption that the inputs are limited to step functions of position and velocity, equation 1 when expressed in terms of the error becomes

$$\mp T = J\frac{d^2E}{dt^2} \tag{2}$$

The solution of equation 2 results in the general phase trajectory equation

$$\left(\frac{dE}{dt}\right)^2 \pm \frac{2T}{J}E = C \tag{3}$$

The ± sign, which determines the direction of motor torque, may be replaced by δ, i.e.

$$\delta = \begin{cases} +1 \text{ for } T<0 \\ -1 \text{ for } T>0 \end{cases} \tag{4}$$

Equation 2 may then be written as

$$\delta = -\frac{J}{T}\frac{d^2E}{dt^2} \tag{5}$$

from which one obtains

$$\left(\frac{dE}{dt}\right)^2 - \delta\frac{2T}{J}E = C \tag{6}$$

Equation 6 can be represented by parabolas 1, 2, 4, shown in Fig. 2(A). At $t=0$, the representative point is at a point in the phase plane corresponding to initial conditions E_0, $\left(\frac{dE}{dt}\right)_0$. Assuming the first delta δ_1, is equal to minus one, the representative point moves along the trajectory 1. If at some time later the delta is changed to plus one, the representative point will follow trajectory 2. The problem is to establish a program for the sign reversal of δ such that the representative point reaches the origin in minimum time.

There is a mathematical theorem[14] which states that for each value of δ there is only one phase trajectory which passes

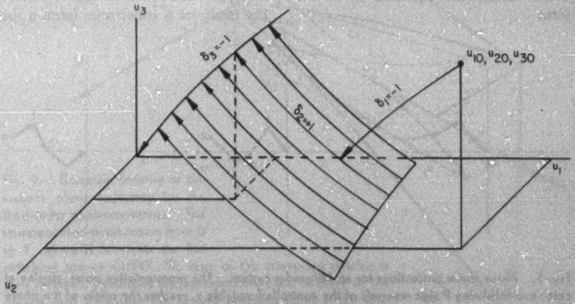


Fig. 3. Phase portrait for a general third-order contactor servo

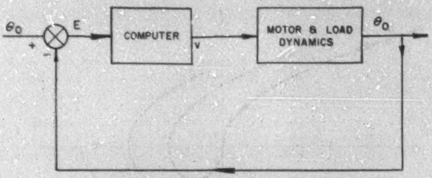


Fig. 4. Block diagram of a general higher order system

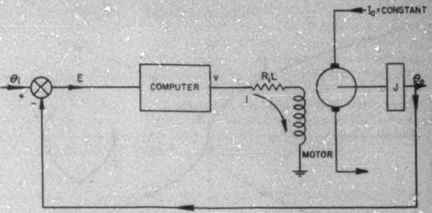


Fig. 6. Third-order contactor servo. The order of the system is increased by accounting for the lag in the motor field

through a given point in the phase plane. Since there exist two possible deltas in the present problem; i.e., $\delta=+1$ and $\delta=-1$, there will be two trajectories which pass through any point, and in particular the origin. These are shown as curves 3 and 4 in Fig. 2(A) and will be referred to as zero trajectories. Every representative point starting at E_0, $\left(\frac{dE}{dt}\right)_0$ to reach the origin, must eventually reach either curve 3 or 4 through some switching sequence. To simplify the analysis, it will be assumed that the initial conditions E_0 and $\left(\frac{dE}{dt}\right)_0$ are in a particular section of the phase plane so that, when proper switching occurs, the representative point reaches the origin along the $\delta=+1$ zero trajectory.

Fig. 2(B) shows the final $\delta_2=+1$ trajectory as 2, a plane curve, extending from the origin. It may be noted that from every point along 2, there radiates a $\delta_1=-1$ trajectory such as 1. The family of these trajectories form a surface whose initial conditions fill the entire upper-half phase plane bounded by curves 2 and 3. For this half of the phase plane the switching criterion would be to let $\delta_1=-1$ until the representative point reaches curve 2, at which time $\delta_2=+1$. The corresponding criterion exists for the other half phase plane where $\delta_1=+1$ and $\delta_2=-1$.

This point of view can be extended to a purely hypothetical third-order system whose differential equation is of the form

$$\delta=F_1\left(E, \frac{dE}{dt}, \frac{d^2E}{dt^2}, \frac{d^3E}{dt^3}\right) \tag{7}$$

and whose general phase trajectory is defined by

$$F_2(\delta, u_1, u_2, u_3)=0 \tag{8}$$

$$F_3(\delta, u_1, u_2, u_3)=0 \tag{9}$$

Equations 8 and 9 form a space curve in the 3-dimensional phase space with coordinates u_1, u_2, and u_3, where

$$u_1=E \tag{10}$$

$$u_2=\frac{dE}{dt} \tag{11}$$

$$u_3=\frac{d^2E}{dt^2} \tag{12}$$

$$\delta=+1 \text{ or } -1 \tag{13}$$

As in the preceding example, there are two trajectories which pass through the origin, one for $\delta=+1$ and one for $\delta=-1$. The discussion is therefore temporarily restricted to initial conditions, u_{10}, u_{20}, u_{30}, in a given half of the phase space such that when proper switching occurs, the representative point reaches the origin along the $\delta=-1$ zero trajectory.

Fig. 3 shows the zero trajectory ($\delta_3=-1$) extending from the origin in the form of a space curve. From every point along the δ_3 curve there radiates a $\delta_2=+1$ curve. The family of δ_2 trajectories forms a surface which terminates in the zero trajectory. The final step is to consider this surface as the termination of a family of trajectories whose initial conditions fill the entire 3-dimensional half-phase space mentioned earlier. The switching criterion for this region is to choose $\delta_1=-1$ until the representative point reaches the switching surface formed by the δ_2 space curves. At the surface, the δ changes sign so that the representative point, traveling in the surface, reaches the zero trajectory and then moves into the origin. As in the previous example, there exists an analogous switching criterion in the other half phase space, where $\delta_1=+1$, $\delta_2=-1$, and $\delta_3=+1$.

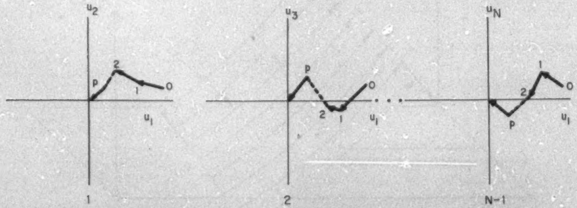


Fig. 5. Phase space projections for an Nth-order system. The representative point, starting at zero, and following P sign reversals of the controlled variable v, reaches the origin of the phase space

The preceding paragraphs represent an effort to create a point of view which can be extended to a phase space of dimensions greater than three. The idea, essentially one of working backwards from the origin of the space, may be summarized for a 3-dimensional space as follows: From the origin there extends the zero trajectory, e.g., $\delta_3=-1$. This path is the termination of a surface composed of a family of trajectories having a driving term δ_2 of opposite sign to that of the zero trajectory; finally this surface is the termination of a family of trajectories whose initial condition fill in the half-phase space.

Mathematical Derivation

It is the purpose of the following section to lend mathematical credence to the material in the previous section. Fig. 4 represents an Nth order system in block form. The motor and load behavior can be defined by the following differential equation

$$-\delta=\sum_{n=0}^{N} b_n \frac{d^n\theta_0}{dt^n} \tag{14}$$

where all system constants are lumped into the b_n coefficients. In terms of the error E, equation 14 becomes

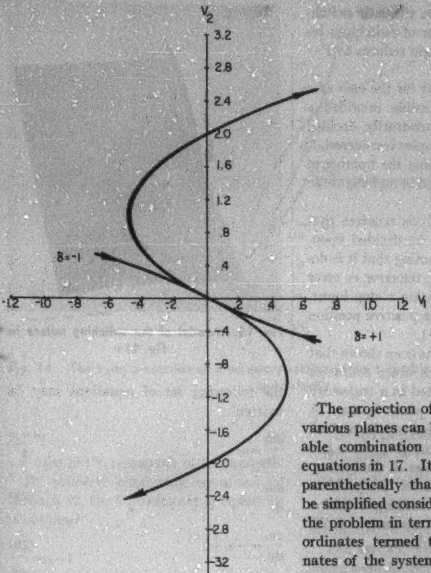


Fig. 7 (left). The v_1-v_2 projection of the zero trajectories for the third-order contactor servo

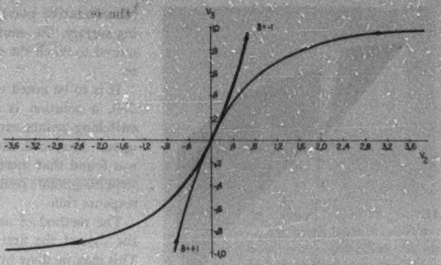


Fig. 8. The v_2-v_3 projection of the zero trajectories for the third-contactor servo

$$\sum_{n=0}^{N} b_n \frac{d^nE}{dt^n} = \delta + \sum_{n=0}^{N} b_n \frac{d^n\theta_i}{dt^n} \qquad (15)$$

an Nth order differential equation in E whose phase space has as co-ordinates

$$u_1 = E$$

$$u_2 = \frac{dE}{dt} \qquad (16)$$

$$\cdots$$

$$u_n = \frac{d^{N-1}E}{dt^{N-1}}$$

Equations 15 and 16 are used to establish a system of differential equations, which, when solved, yield the trajectory of the representative point

$$\frac{du_1}{dt} = u_2$$

$$\frac{du_2}{dt} = u_3$$

$$\cdots \qquad (17)$$

$$\frac{du_{N-1}}{dt} = u_N$$

$$\frac{du_N}{dt} = \sum_{n=0}^{N-1} -\frac{b_n}{b_N}\, u_{n+1} + \frac{\delta}{b_N} + \sum_{n=0}^{N} \frac{b_n}{b_N} \frac{d^n\theta_i}{dt^n}$$

The projection of the trajectory in the various planes can be obtained by a suitable combination and solution of the equations in 17. It should be noted here parenthetically that such a solution can be simplified considerably by considering the problem in terms of a new set of co-ordinates termed the principal co-ordinates of the system, obtained through a linear homogeneous transformation. Since the transformation is not pertinent to this discussion, it will be assumed that the appropriate equations have been solved for the N-1 projections of the general trajectory in all the planes common to the u_1 axis, as shown in Fig. 5.

$$\begin{aligned} &h_1(u_1, u_2, \delta) = 0 \\ &h_2(u_1, u_3, \delta) = 0 \\ &\cdots\cdots \\ &h_{N-1}(u_1, u_N, \delta) = 0 \end{aligned} \qquad (18)$$

The object is to determine the loci of points at which δ changes sign, such that the final trajectory passes through the origin. If points 1, 2 . . . P, represent the switching points, let $\delta = \delta_1$; e.g., $\delta_1 = -1$, between 0 and 1, $\delta_2 = -\delta_1$ between 1 and 2, etc. The problem then resolves itself into the solution of equation 18 (with the proper δ's inserted) for the unknown co-ordinates of the P points.

The following is noted from Fig. 5:

N = order of the system
$N-1$ = number of independent phase planes
P = number of switchings
$P+1$ = number of equations per plane

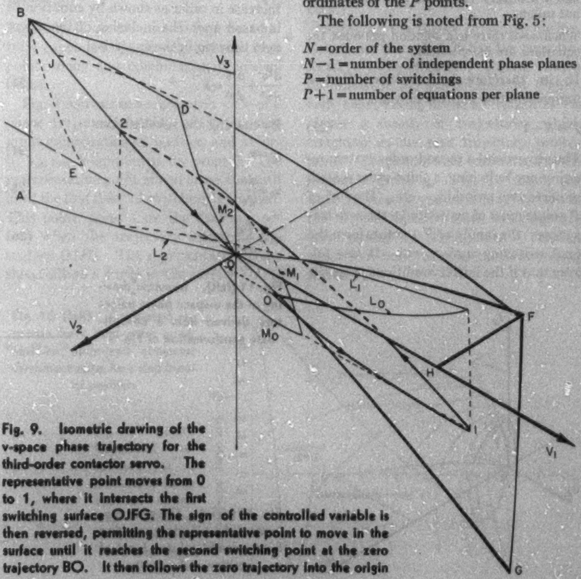


Fig. 9. Isometric drawing of the v-space phase trajectory for the third-order contactor servo. The representative point moves from 0 to 1, where it intersects the first switching surface OJFG. The sign of the controlled variable is then reversed, permitting the representative point to move in the surface until it reaches the second switching point at the zero trajectory BO. It then follows the zero trajectory into the origin

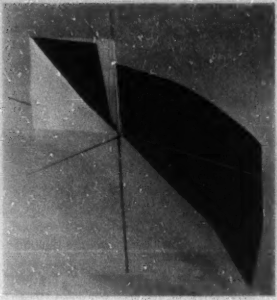


Fig. 10. Model of the switching surface in Fig. 9

$(P+1)(N-1)$ = total number of equations
NP = number of unknowns

Let a be defined by

$$NP-(P+1)(N-1)=a \tag{19}$$

$$(\text{variables})-(\text{equations})=a \tag{20}$$

or

$$P=N-1+a \tag{21}$$

The following can be stated for possible values of a:

When $a<0$, assuming the equations are independent, there is no solution.

When $a>0$, assuming there is a solution, it has a arbitrary constants and therefore is not unique.

When $a=0$ there is a solution provided the equations are independent.

It can therefore be concluded that a unique solution is possible when

$$P=N-1 \tag{22}$$

Thus in general a second-order system requires one switching, a third-order system requires two switchings, etc. The loci of P points must of necessity be the zero trajectory; the family of P-1 points form the final switching surface, etc. It also follows that if the initial conditions place the representative point on the mth switching surface, the number of switchings required to reach the origin reduces to $P-m$.

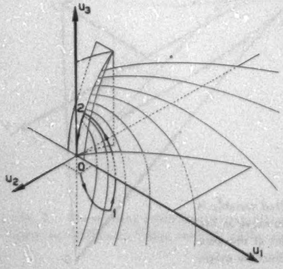


Fig. 11 (left). Isometric drawing of the u-space phase trajectory derived from a co-ordinate transformation of Fig. 9

It is to be noted that for the case of $a>0$, a solution is possible provided a switching points are arbitrarily decided upon. In specific examples considered, it was found that increasing the number of switching points resulted in an increase in response time.

The method of analysis requires that the sign of the first δ be decided upon. This may be done by noting that it is desirable to begin reducing the error or error rate immediately following a step input. On this basis, for either positive position or velocity input $\delta_1=-1$.

In summarizing, it has been shown that the behavior of an Nth order contactor servo may be represented as a trajectory in an N-dimensional phase space. A unique switching criterion can be derived when the number of switching points P is one less than N, the order of the system. For given initial conditions, the problem of establishing the switching criterion then resolves itself into a solution of NP equations in as many unknowns.

Application of Phase Space Analysis

The method proposed in the previous section is applied to the third-order on-off servomechanism shown in Fig. 6. The increase in order as shown by equation 23 is based upon the inclusion of the motor field time lag in the analysis.

$$\frac{d^3\epsilon}{d\tau^3}+\frac{d^2\epsilon}{d\tau^2}=\delta \tag{23}$$

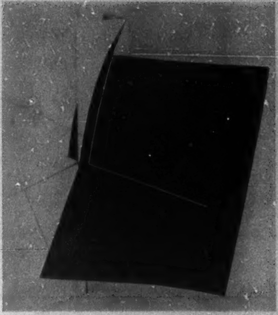


Fig. 12. Model of the switching surface in Fig. 11

By making the substitutions

$$u_1=\epsilon \tag{24}$$

$$u_2=\frac{d\epsilon}{d\tau} \tag{25}$$

$$u_3=\frac{d^2\epsilon}{d\tau^2} \tag{26}$$

the following set of equations may be written

$$\frac{du_1}{d\tau}=u_2 \tag{27}$$

$$\frac{du_2}{d\tau}=u_3 \tag{28}$$

$$\frac{du_3}{d\tau}=-u_3+\delta \tag{29}$$

These equations represent the differential equations which must be solved to yield the general trajectory. To simplify the mathematics the principal co-ordinates defined by

$$v_1=u_1-u_3 \tag{30}$$

$$v_2=u_2+u_3 \tag{31}$$

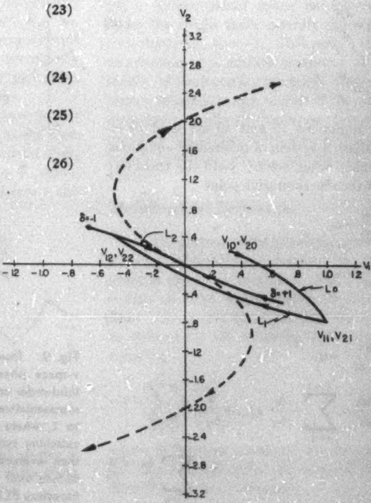


Fig. 13 (right). The v_1-v_2 projection of a trajectory resulting from a combined step input of position and velocity

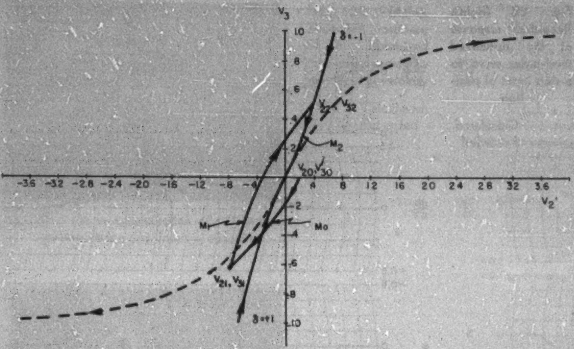


Fig. 14. The v_2-v_3 projection of a trajectory resulting from a combined step input of position and velocity

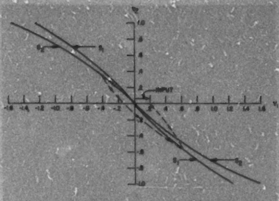


Fig. 15. The v_1-v_2 projection of the switching curves for a step input of position. The representative point projection travels along the dashed line

$$v_3 = u_3 \tag{32}$$

are used in the remainder of the analysis.

By properly combining equations 27 through 32, the following set of equations is obtained

$$\frac{dv_1}{d\tau} = v_2 - \delta \tag{33}$$

$$\frac{dv_2}{d\tau} = \delta \tag{34}$$

$$\frac{dv_3}{d\tau} = -v_3 + \delta \tag{35}$$

The simultaneous solution of equations 33, 34, and 35 yields

$$2v_1 - \delta(v_2 - \delta)^2 = C_1 \tag{36}$$

and

$$v_2 + \delta \ln(1 - \delta v_3) = C_2 \tag{37}$$

which are the projections of the phase trajectory in the v space. The initial conditions v_{10}, v_{20}, v_{30} serve to determine C_1 and C_2. It should be noted that a change in initial conditions results only in the translation of the trajectory projections in the v_1-v_2 and v_2-v_3 planes.

The trajectories passing through the origin, given by equations 38 and 39

$$(2v_1 + \delta) - \delta(v_2 - \delta)^2 = 0 \tag{38}$$

$$v_2 + \delta \ln(1 - \delta v_3) = 0 \tag{39}$$

are plotted in Figs. 7 and 8.

The arrows, which indicate increasing time, were determined by solving equation 34.

$$v_2 - v_{20} = \delta\tau \tag{40}$$

If Δ is defined as an increment, then

$$\Delta v_2 = \delta(\Delta\tau) \tag{41}$$

Equation 41 enables one to find the total time expended in moving between two points along any combination of trajectories.

Fig. 9 represents an isometric view of a phase trajectory having an initial step input combination of position and velocity v_{10}, v_{20}, designated as the point 0. The representative point moves along the path until the first point of switching is reached. This point, point 1 on the drawing, occurs when the trajectory intersects the surface *OJFG*. The representative point then follows a curve in the surface *OJFG* until it reaches the second switching point 2, which is the intersection of the trajectory 1-2 and the zero trajectory *BO*. It then follows *BO* into the origin. The curves L_0, L_1, L_2, are projections of the trajectory in the plane of axes v_1, v_2, and curves M_0, M_1, M_2 are projections of the trajectory in the v_2, v_3 plane.

The final switching curve *BO* is represented as the intersection of two surfaces *OABC* and *OEBD*. Curve *OA* corresponds to the curve $\delta = -1$ in Fig. 7, and *OABC* is the surface through *OA* with vertical elements parallel to the v_3 axis. Curve *OD* corresponds to $\delta = -1$ in Fig. 8, and *OEBD* is the surface through *OD* with horizontal elements parallel to the v_1 axis. Curve *OF* is the intersection of the v_1-v_2 plane and the switching surface *OJFG*.

Applying the theory presented earlier, it may be seen that *BO* is the zero trajectory. The switching surface which terminates in *BO* was obtained by graphically constructing in the v_1-v_2 and v_2-v_3 planes a family of trajectories which terminate on the zero trajectory projections. The results were transferred to the isometric drawing. Fig. 11 shows the same trajectory and switching surface in the u-phase space. Both drawings are accompanied by photographs of models which were built to represent the switching surface.

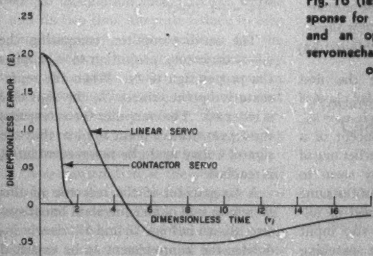


Fig. 16 (left). Computed response for a third-order linear and an optimized contactor servomechanism for a step input of position

Fig. 17 (right). Computed response for a third-order linear and an optimized contactor servomechanism for a step input of velocity

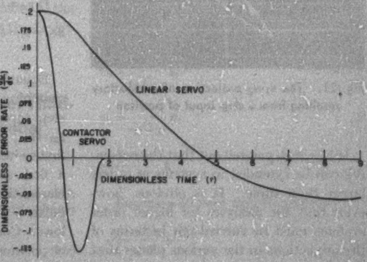


Fig. 18. The v_1-v_2 projection of the zero trajectories

Fig. 19. The v_2-v_3 projection of the zero trajectories

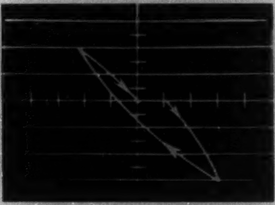


Fig. 20. The v_1-v_2 projection of a trajectory resulting from a step input of position

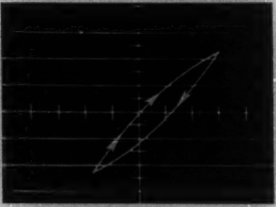


Fig. 21. The v_1-v_3 projection of a trajectory resulting from a step input of position

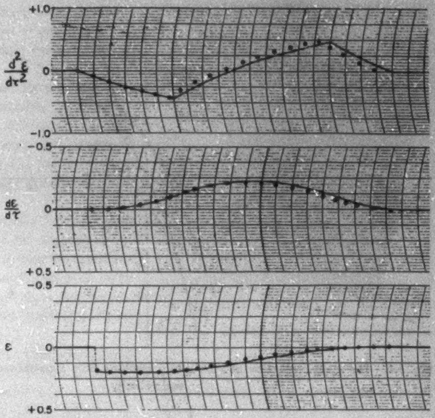


Fig. 22 (right). Recorded response of the optimized third-order servo to a step input of position

. . . . Calculated
——— Recorded

It has been pointed out that the reasoning can be extended to systems of orders larger than three. It is obvious, however, that the analyses for higher order systems must be carried out in terms of the projections in the various planes that make up the N-dimensional space. The remainder of the analysis will therefore be devoted to obtaining the switching criteria in terms of one particular plane, in this case the v_1-v_2 plane. Where necessary, the plane-switching criterion will be correlated with the surface and space-curve-switching criterion.

Figs. 13 and 14, which display the trajectory projections, were derived on the following basis: As a third-order system, two switchings are required. The final path (also the final switching curve) is the zero trajectory, and because the input is positive the initial path must be along a path corresponding to $\delta_1=-1$. The problem remaining, to determine the locus of v_{11}, v_{21}, is solved as follows. Equations 36 and 37 are used to write six equations corresponding to the L and M projections. The simultaneous solution of these equations yields

$$2v_{11}-\delta_1\left\{1+2\ln^2\left[1+\sqrt{1-e^{-\delta_1 v_{21}}(2-e^{-\delta_1(v_{21}-v_{20})})}\right]\right\}+\delta_1(v_{31}+\delta_1)^2=0 \quad (42)$$

If $v_{20}=0$, equation 42 becomes

$$2v_{11}-\delta_1-\delta_1 2\ln^2(2-e^{-\delta_1 v_{21}})+\delta_1(v_{21}-\delta_1)^2=0 \quad (43)$$

Equations 42 and 43 are the first switching curves for a particular class of step inputs, namely for which $v_{10}=u_{10}$; $v_{20}=u_{20}$; and $v_{30}=0$. The concept of a first switching curve and the earlier one of a first switching surface may seem to clash; they are, however, part of the same pattern which may be explained as follows: Consider only the class step input of position. The locus of first switching points forms a space curve in the switching surface whose projection in the v_1-v_2 plane is given by equation 43. Similarly for a class of step inputs of position plus velocity, where the velocity is specified, the locus of first switching points forms a space curve in the switching surface, which when projected in the v_1-v_2 plane becomes the curve given by equation 42. Note that there is a different projection for each value of v_{20}.

The second sign reversal curve may be obtained from equation 38, the zero trajectory projection in the v_1-v_2 plane.

$$2v_1+\delta_3-\delta_3(v_2-\delta_3)^2=0 \quad (44)$$

These sign reversal curves, plotted in Fig. 15 for $v_{20}=0$, were obtained as follows

$$S_1(v_1, v_2)=\begin{cases}\text{equation 43 } \delta_1=-1 \ v_{21}<0\\ \text{equation 43 } \delta_1=+1 \ v_{21}>0\end{cases} \quad (45)$$

$$S_2(v_1, v_2)=\begin{cases}\text{equation 44 } \delta_2=-1 \ v_2>0\\ \text{equation 44 } \delta_2=+1 \ v_2<0\end{cases} \quad (46)$$

The switching sequence is then

$$1\begin{cases}\delta_1=-1 & S_1>0\\ \delta_1=+1 & S_1<0\end{cases} \quad (47)$$

$$2\begin{cases}\delta_2=-1 & S_2>0\\ \delta_2=+1 & S_2<0\end{cases} \quad (48)$$

The servo computer, comparing the v_1-v_2 trajectory projection to S_1 applies the proper sign to δ_1. When the representative point crosses S_1, the sign of δ is reversed. The computer then compares the representative point to S_2 so that the sign of δ may again be reversed when S_2 is reached.

A comparison of the response of this system to that of an equivalent linear system, shown in Figs. 16 and 17, clearly indicates the improvement to be expected

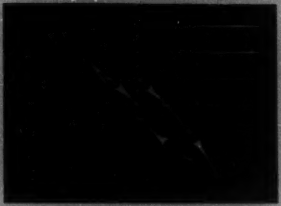


Fig. 23. The v_1-v_2 projection of a trajectory resulting from a step input of velocity

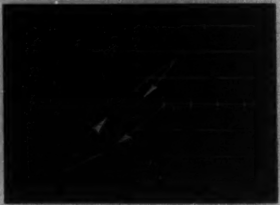


Fig. 24. The v_2-v_3 projection of a trajectory resulting from a step input of velocity

from the use of an optimum-switching on-off servo.

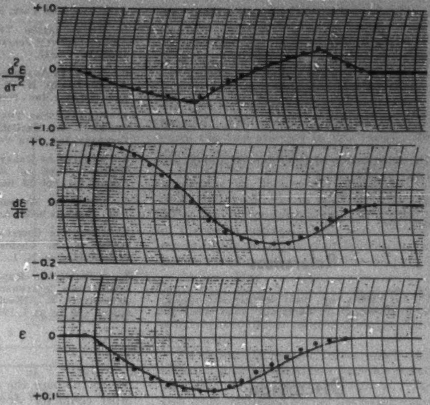
Computer Study

Figs. 18 through 25 represent results obtained from an analogue computer study of a third-order on-off servo in which the foregoing switching criterion was employed. In all cases data taken from the photographs were found to agree with calculated behavior to within experimental accuracy.

An examination of Fig. 22 leads, in this particular example, to what at first appears to be a new interpretation of the concept of a unique switching criterion. The curve of error acceleration may be viewed as one of torque build-up. It can therefore be seen that the derived switching criterion is essentially a program of torque modulation such that the torque and its two time integrals reduce to zero simultaneously. The condition for the case of velocity input, as demonstrated by Fig. 25 is quite similar. The torque in this case must be manipulated to permit the error rate to overshoot, since the error, the time integral of error rate, begins and must eventually terminate at zero.

A re-examination of the switching criterion for the second-order on-off servo leads to similar results, the major differences being that for the second-order system torque build-up is assumed instantaneous while for the third-order system it is exponential.

Fig. 25 (right). Recorded response of the optimized third-order servo to a step input of velocity



Summary Conclusions

The purpose of this paper is to describe a method for obtaining the switching criterion for an *N*th order on-off servo which reduces system error and derivatives of error to zero in a minimum time.

System behavior is described in the principal co-ordinate phase space. A switching criterion is then established from the simultaneous solution of a number of equations representing trajectory projections; the first trajectory passing through a point defined by the initial conditions, and the final one passing through the origin.

The mathematical analysis employed imposes two restrictions on the results: 1. inputs are limited to step-functions; 2. the system must be describable over every interval of its operation by a linear differential equation.

Most of the data presented were obtained from an analogue computer study. In view of the results, it is felt that the method merits further development to the point where it forms a basis for the design of on-off control systems.

Appendix I. Derivation of Third-Order System Differential Equations

The following equations describe each section of the third-order on-off servo shown in Fig. 6.

$$v=-\delta V \tag{49}$$

$$v=Ri+L\frac{di}{dt} \tag{50}$$

$$\phi=K_\phi i \tag{51}$$

$$T=K_2\phi \tag{52}$$

$$T=J\frac{d^2\theta_0}{dt^2} \tag{53}$$

V is the magnitude of the computer output voltage; δ is a function of the error, error rate, and error acceleration, having a value of +1 or −1, as determined by the computer.

The combination of equations 49 through 52 results in

$$-\delta V\frac{K_2K_\phi}{R}=T+\tau_\phi\frac{dT}{dt} \tag{54}$$

and the combination of equations 53 and 54 yields

$$-\delta V\frac{K_2K_\phi}{RJ}=\tau_\phi\frac{d^3\theta_0}{dt^3}+\frac{d^2\theta_0}{dt^2} \tag{55}$$

It will be assumed that *V* is large enough to drive the motor to torque saturation T_M. From equation 54, therefore

$$T_M=V\frac{K_2K_\phi}{R} \tag{56}$$

New dimensionless variables may be defined as follows

$$\tau=\frac{t}{\tau_\phi} \tag{57}$$

$$\Theta_i=\frac{J}{\tau_\phi^2T_M}\theta_i \tag{58}$$

$$\Theta_0=\frac{J}{\tau_\phi^2T_M}\theta_0 \tag{59}$$

$$\epsilon=\frac{J}{\tau_\phi^2T_M}\mathrm{E} \tag{60}$$

Equation 55 may then be written

$$-\delta=\frac{d^3\Theta_0}{d\tau^3}+\frac{d^2\Theta_0}{d\tau^2} \tag{61}$$

The dependent variables are related by

$$\frac{d^n\Theta_0}{d\tau^n}=\frac{d^n\Theta_i}{d\tau^n}-\frac{d^n\epsilon}{d\tau^n}\quad n=0,1,2\ldots \tag{62}$$

Under the assumption that

$$\frac{d^n\Theta_i}{d\tau^n}=0\quad n\geqq 2 \tag{63}$$

the substitution of equations 62 and 63 into equation 64 yields

$$\frac{d^3\epsilon}{d\tau^3}+\frac{d^2\epsilon}{d\tau^2}=\delta \tag{64}$$

References

1. Principles of Servomechanisms (book), **G. S. Brown, D. P. Campbell.** John Wiley and Sons, Inc., New York, N. Y., 1948.

2. Regeneration Theory, **H. Nyquist.** *Bell System Technical Journal*, New York, N. Y., Jan. 1932, pp. 126–47.

3. Network Analysis and Feedback Amplifier Design (book), **H. W. Bode.** D. Van Nostrand and Company, Inc., New York, N. Y., 1945.

4. Fundamental Theory of Servomechanisms (book), **L. A. MacColl.** D. Van Nostrand and Company, Inc., New York, N. Y., 1945.

5. Analysis of Relay Servomechanisms, **H. K. Weiss.** *Journal of the Aeronautical Sciences*, Easton, Pa., vol. 13, no. 7, July 1946, pp. 364–76.

6. Nonlinear Mechanics (book), **N. Minorsky.** J. W. Edwards Bros., Ann Arbor, Mich., 1947.

7. Theory of Oscillations (book), **A. A. Andronow, C. E. Chaiken.** Princeton University Press, Princeton, N. J., 1949.

8. A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms, **R. J. Kochenburger.** *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

9. Sinusoidal Analysis of Feedback-Control Systems Containing Nonlinear Elements, **E. Calvin Johnson.** *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 169–81.

10. Optimization of Nonlinear Control Systems by Means of Nonlinear Feedbacks, **R. S. Neiswander, R. H. MacNeal.** *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 262–72.

11. A Phase-Plane Approach to the Compensation of Saturating Servomechanisms, **Arthur M. Hopkin.** *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 631–39.

12. Nonlinear Techniques for Improving Servo Performance, **D. McDonald,** *Proceedings*, National Electronics Conference, Inc., Chicago, Ill., vol. 6, pp. 400–21.

13. The Mathematics of Physics and Chemistry (book), **H. Marganau, G. M. Murphy.** D. Van Nostrand and Company, Inc., New York, N. Y., 1948.

14. Theory of Functions of a Real Variable (book), **L. M. Graves.** McGraw-Hill Book Company, Inc., New York, N. Y., 1946, chap. 9.

◆

Discussion

A. Lange (Massachusetts Institute of Technology, Cambridge, Mass.): The authors are to be congratulated on their interesting paper. Contactor servo design and analytic techniques for nonlinear systems are subjects of great importance to the systems engineer. A paper such as this, making new contributions to both subjects, is especially welcome.

It should be remarked at the outset that nonlinear equations with discontinuous characteristics offer special problems to the analyst. In dealing with such problems as contactor servo design and mechanical hysteresis,[1] the engineer is forced to proceed with very little assistance from the mathematician. Most of the well-known theorems and indexes[2,3] which have proven so useful in the analysis of nonlinear systems with continuous characteristics have no meaning for discontinuous systems.

It is therefore not clear that existence theorems[4] are of any particular value, or that they are even valid. For example, in the second-order system discussed in this paper, near the origin, the representative point may be governed by any one of four different nonlinear equations.

Actually, the statement of the problem[5] (which determines the switching criteria) provides a more definitive, and thus restrictive, condition of uniqueness than a Lipschitz condition does. Since the problem statement is closely related to the physics of the problem, the physical meaning is clear and there is no doubt that when a switching line is reached cf. line 2, Fig. 2(B), the representative point will move toward, not away from, the origin.

Another mathematical scheme which did not appear sufficiently clear to me was that of the principle co-ordinate phase space. The transformation of co-ordinates is often an expedient mathematical device. However, by making known the motivation of advantage of the transformation, the loss of physical meaning which accompanies it is easier to accept.

Although the analysis here requires that the forcing function be restricted to a step command of position or velocity, it is not clear that these are the only allowable inputs. In fact, it would have been interesting to compare the response of the third-order contactor servo and the equivalent linear servo to other inputs, in particular, sinusoidal and step inputs of different magnitudes. It would be desirable to know in what sense the linear system is equivalent. The restriction on piecewise linearity also deserves some amplification.

It is interesting to note the use of an analogue computer in this study. Analogue computers, especially the high-speed machines, are powerful tools for the nonlinear analyst.[6] At the instrumentation laboratory, the general-purpose simulator has been used to examine the phase plane in-the-large, to determine the location of stable and unstable limit cycles, separtrices, hard and soft operating conditions, and other such phenomena. It might be useful to evaluate higher order systems on the phase plane by using the analogue computer, especially to determine the variation of the phase trajectories with the important system parameters.

References

1. The Dynamics of Automatic Controls (book), **R. C. Oldenbourg, H. Sartorious** (Translated by **H. L. Mason**). American Society of Mechanical Engineers, New York, N. Y., 1948, pp. 159ff.

2. See reference 6 of the paper, pp. 62ff.; pp. 105ff.

3. See reference 7 of the paper, pp. 118ff.

4. Lectures on Differential Equations (book), **Solomon Lefschetz.** Princeton University Press, Princton, N. J., 1946, pp. 23ff.

5. See reference 12 of the paper.

6. A Topological and Analog Computer Study of Certain Servos Employing Non-Linear Electronic Components, **R. C. Lathrop.** *Ph. D. Dissertation*, University of Wisconsin, Madison, Wis., 1951.

Rufus Oldenburger (Woodward Governor Company, Rockford, Ill.): In 1944, I obtained the optimum solutions for the second- and third-order cases exemplified by equations 2 and 23. Subsequent development of the theory and efforts to apply it show that noise places a severe limitation on its use. The error E is made up of signal and noise. One wishes to respond to the signal but not to the noise. It is in this sense that I refer to noise.

Consider the quantity $\dot{E}^2$ that occurs in equation 6. To simplify the discussion, suppose that the signal is zero and that E is made up only of the noise $E_0 \sin \omega t$. It will be recalled that ω is 2π times the frequency. The first derivative $\dot{E}$ is now $\omega E_0 \cos \omega t$. It follows that

$$\dot{E}^2=\omega^2E_0^2\cos^2\omega t \tag{65}$$

For the absolute value $|\dot{E}|$ of $\dot{E}$ the quantity $|\dot{E}|\dot{E}$ is now a periodic function with the same period as $\sin \omega t$ but amplitude $\omega^2E_0^2$. The second derivative $\ddot{E}$ is $-\omega^2E_0 \sin \omega t$, and thus has the amplitude ω^2E_0. As far as frequency is concerned the quantity $\dot{E}^2$, or $|\dot{E}|\dot{E}$, behaves like the second derivative $\ddot{E}$. The amplitude of the noise in $\dot{E}^2$ goes up as the square of the frequency. The amplitude of the noise in $\dot{E}^2$, or $|\dot{E}|\dot{E}$, also goes up as the square of the amplitude E_0. It has been my experience that a term $\dot{E}^2$ or $|\dot{E}|\dot{E}$ gives rise to a noise considerably worse than that in $\dot{E}$, and on the order of the noise in $\ddot{E}$. Because of the magnitude of the noise I have found that virtually identical transients can be obtained in many cases if the quantity $|\dot{E}|$ in $|\dot{E}|\dot{E}$ is replaced by a sort of "average" value that $|\dot{E}|$ takes on during the transients. We have found that one half of the maximum value of $|\dot{E}|$ is a good choice for the average value. The controlling function is then of the form $E+\alpha\dot{E}$ for a constant α, and is thus a linear function. The advantages of using a linear function in place of the quadratic in equation 6 are obvious.

For the third-order case, equation 23, the controlling functions in equations 42 and 43 involve the square $\ddot{E}^2$ of the second derivative of E. With $E=E_0 \sin \omega t$ we have

$$\ddot{E}^2=\omega^4E_0^2\sin^2\omega t \tag{66}$$

Now, the fourth derivative of E is $\omega^4E_0 \sin \omega t$. As far as frequency is concerned, the quantity $\ddot{E}^2$ thus behaves like a fourth derivative. The amplitude of the noise in $\ddot{E}^2$ goes up as the fourth power of the frequency.

If we consider higher order systems, the situation is much worse. As anyone who has worked with physical differentiators knows, the problems involved even in the engineering use of a second derivative are great. Because noise is inherently present in all physical problems it appears to me that the development of the theory under discussion for higher order systems if of little or no practical importance (although I, myself, have actually worked along these lines).

I believe that a combination of statistical theory (shades of Norbert Wiener) and nonlinear theory is indicated in many switching problems.

I. Bogner and L. F. Kazda: In response to questions raised by Mr. Lange, it may be said that, to simplify the analysis, the authors deliberately limit it to a system which is linear over every interval of its operation. Between control variable reversals, therefore, the servo is describable by a linear differential equation with real constant coefficients. Only the sign of the control variable is assumed nonlinear. It is therefore considered that "piecewise linear" rather than "nonlinear" represents an accurate description of the servo.

Because discussion is restricted to phase-space trajectories describing a piecewise or sectionally linear servo, the existence theorem which is employed (and is applicable in linear theory) is also applicable to each trajectory section.

The ease with which one may describe servo behavior, and the resulting simplification in deriving a switching criterion are the primary reasons for transforming to the principal co-ordinates. In addition, one finds, for example, that in particular principal co-ordinate planes, trajectory projections are identical in shape, a change in initial conditions resulting simply in a translation. This facilitates a graphical rather than an analytic derivation of the switching curves. Another factor of importance is that in a particular problem a trajectory increment, projected along one of the principal co-ordinates, may be directly proportional to time. This enables one to obtain the response of the servo as a function of time directly.

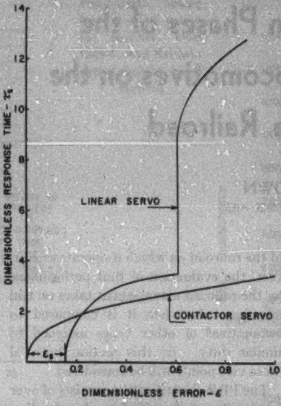


Fig. 26. Response time for third-order linear and contactor servomechanisms with no output damping for step-input of position

The question of inputs other than step functions was not considered. The authors agree that a sinusoidal response may be interesting, but question its quantitative significance in connection with control of the type discussed. Under another phase of the project, a study was carried out to determine the effects of two step inputs, one following the other before steady state is reached. Under these conditions the optimumly switched servo has a response time on the order of one-eighth to one-tenth that of the equivalent linear system.

The criteria used to compare the linear and contactor servomechanism were omitted by oversight. The equivalent linear system was composed of the same motor and load as the contactor servo, and utilized ideal proportional-plus-error rate control so as to give a 25 per cent first overshoot. The settling time in a linear system is defined as the time necessary for the error to reach and remain at a point less than a certain percentage of its initial value. The settling time for a linear servo is therefore independent of the size of the initial error. The corresponding time for the equivalent contactor servo, on the other hand, increases as a function of the error. A comparison based on this type of definition would therefore show the contactor servo to be better for small errors and correspondingly worse for large error signals. A more realistic comparison results if the response is defined in terms of the time necessary to reach a fixed error. Thus for the linear system the response time τ_s is taken as the time necessary to reach and remain less than a fixed small error ϵ_s. The response time for the contactor servo is the time necessary to the error to reduce to zero. The accompanying figure compares the two systems in this manner for various size-step inputs.

The authors agree with Dr. Oldenburger concerning the limitations with regard to noise, and would like to point out that this is not the only problem taxing the ingenuity of the servo engineer. Inaccurate switching resulting from practical limitations, or inherent nonlinearities pose other interesting problems. However, limitations such as these have been resolved by applying the dual mode technique which makes use of a small region of linear control about the phase space origin.

To answer the specific points which were raised, it may be significant that there exists an alternative to operating with $\ddot{E}$. The two space-switching curves discussed under equation 42 can be projected in the u_1-u_2 or error-error rate plane instead of the v_1-v_2 plane. This admittedly does not eliminate the noise problem; it does, however, reduce the problem by eliminating consideration of all but the first derivative term.

In this respect the authors wish to point out that under another phase of the contract, considerable success was achieved in nonlinear switching in the $E-\dot{E}$ plane, in spite of noise considerations. The servo response was markedly superior to response resulting from controlling functions of the $(E+\alpha\dot{E})$ type.

It may be of interest that in addition to the experimental and analytical verification of this work which has been obtained by the authors of this paper, further supporting evidence is presented.[1-4]

REFERENCES

1. See reference 10 of the paper.

2. See reference 11 of the paper.

3. AUTOMATIC AND MANUAL CONTROL (book), A. Tustin. Academic Press, Inc., New York, N. Y., 1952, pp. 295–99.

4. NONLINEAR OPTIMIZATION OF RELAY SERVOS, L. M. Silva. *Thesis*, University of California, Berkeley, Calif., 1953.

Some Application Phases of the Ignitron Rectifier Locomotives on the Pennsylvania Railroad

F. D. BROWN
ASSOCIATE MEMBER AIEE

WHEN two ignitron rectifier freight locomotives (class *E2C* and *E3B*) were first put into service on the Pennsylvania Railroad (PRR) approximately 2 years ago, the event could well have been taken as marking the golden jubilee of main-line electrification in this country. Certainly their successful design and operation are a fitting tribute to those who in 1902 first took a contract to equip a high-speed railroad line with an electrification using single-phase series-type railway motors.

It was recognized then, and always has been, that the single-phase commutator-type railway motor is not in itself as economical or efficient as its d-c competitor. However, opposed to this is placed the simplification and economy of the transmission system, together with more economical speed control, and the omission of rheostatic heat losses.

Now, the rectifier locomotive parts company with the a-c motor. Curiously enough, it was largely through the persistent efforts of those men who overcame the heavy prejudice for direct current, and proved the single-phase motor a success, that the last half-century has seen the widespread development of an a-c electrification which, in turn, puts to such excellent use the advantages of rectifier motive power.

The locomotives themselves have previously been well described[1] by Whittaker and Hutchison. This paper will present some of the more interesting aspects of their application based on the first 2 years of test and revenue service. Principal considerations will be their traction performance in comparison with other locomotives, and the subject of telephone influence.

Performance

FREIGHT TRAFFIC PATTERN

To present a realistic indication of the performance of any locomotive, it is necessary to understand the geographical characteristics and usual traffic pattern of the railroad on which it operates. Further, the evaluation of that performance by the railroad management takes on real meaning only when it is compared to locomotives of other types assigned to similar duty. In this section both of these viewpoints will be considered.

The PRR electrification consists of over 2,200 miles of electrified track extending from New York to Washington, D.C., and to Harrisburg, Pa. From Enola, the huge freight classification yard across the river from Harrisburg, coal and mineral trains are operated to terminals in Jersey City, South Amboy, Trenton, Philadelphia and Baltimore. The ore trains are hauled from Baltimore to Enola, and in the future will probably be hauled from South Philadelphia to other inland terminals. Fast merchandise trains are operated over all these routes and also between the various terminals from Washington to Jersey City.

Arranged weekday electric freight service amounts to 104 trains which average about 30 miles per hour on main-line tracks. Extra trains are also operated when required for special movements. While distances between electrified freight terminals vary only from 25 to 228 miles, analysis of movement shows that the 104 freight trains accumulate approximately 12,300 miles each weekday.

Merchandise trains often consist of 125 cars, totaling 6,250 tons. Coal or mineral trains may be made up of 150 to 160 cars, totaling 13,500 tons. Westward trains, which normally contain a large percentage of empty cars, may have as many as 175 cars, and 5,250 tons. Long, heavy trains, some exceeding 1.5 miles in length, are hauled by new 2-unit electric locomotives or by the older electric locomotives operating in multiples of two or three.

Practically all of the through passenger and freight traffic in electrified territory is hauled by 241 electric road locomotives, most of which can be operated either in passenger or freight service, depending upon traffic demands. A total of 97 are actually geared for freight service. Generally speaking, the freight trains are hauled by the type *P-5a* locomotive, and passenger trains by the *GG-1*.

The tonnage ratings of various PRR locomotives over the ruling grade line between Columbia and Parkesburg, Pa., are given in Table I. From this it follows that it takes three *GG1* or three *P5A* locomotives to match the rating of one 2-unit rectifier locomotive.

PERFORMANCE COMPARISON WITH OTHER LOCOMOTIVES

In making any locomotive comparison, it should be remembered that the maximum tractive force which can be transmitted is limited by two mechanical factors. The first is the total weight on those wheels which actually transmit tractive force, which is fixed for a particular locomotive. The second is the adhesion which can be maintained between the surfaces of the driving wheels and track rails. The amount of propelling force which can be applied to the rims at the driving wheels is obviously limited by the point at which slipping begins. The degree of adhesion which can be maintained decreases as speed increases. It is also appreciably reduced when the running rails are wet.

The 2-unit rectifier locomotive will be compared with two *GG-1* locomotives coupled together and geared for 90 miles per hour (mph). It will also be compared with a 3-unit 4,800-horsepower (hp) diesel-electric locomotive, which has approximately the same weight on the driving wheels and which also has the same number and type of traction motors as the rectifier locomotive. It will be apparent that, with an essentially unlimited source of electric power, the rectifier locomotive does much more with these motors than does the diesel locomotive with its fixed hp engines. Actually, the full possible performance of the rectifier locomotive will not be indicated since certain current restrictions have arbitrarily been kept on

Table I. Tonnage Ratings

Locomotive Type	Flat Tons	Weight on Drivers, Pounds
1-unit *P5A*	4,350	225,000
2-unit *E2B*	7,800	480,000
1-unit *GG1*	4,250	300,000
2-unit rectifier locomotive	12,000	750,000

Paper 54-116, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing November 23, 1953.

F. D. BROWN is with the Westinghouse Electric Corporation, Philadelphia, Pa.

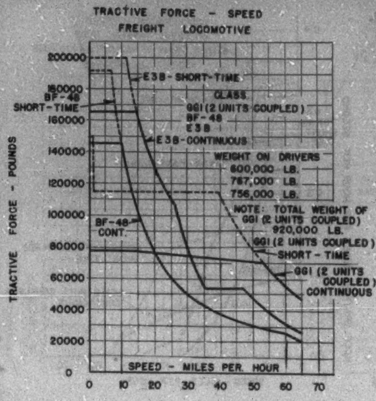


Fig. 1 (left). Speed tractive-effort curves for rectifier, class GG-1, and Baldwin diesel freight locomotives

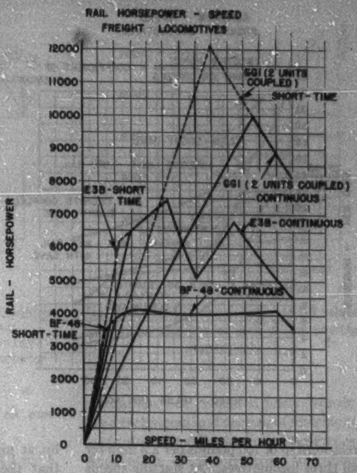


Fig. 2 (right). Speed rail-hp curves

it to date as a deliberate effort to make all such evaluations conservative.

Fig. 1 shows the tractive effort versus speed for each of the three types of locomotives. Note that the *GG-1*'s, designed for high-speed passenger service, do not have a high tractive effort at starting. They can begin with 150,000 pounds (25 per cent of weight on drivers) but, after starting, must reduce to 110,000 pounds to accelerate the train. If required to haul a train up a long grade, they would have to reduce further to approximately 75,000 pounds up to a speed of 53 mph.

Because of differences in weight, gearing, and wheel size, the rectifier locomotive can provide a starting tractive effort of 200,000 pounds while using only the same motor currents that produce 192,000 pounds from the diesel. The continuous tractive force for the rectifier locomotive is purposely limited by "notching restrictions" which allow the engineman to pull only certain values of traction motor current for each master controller position. This results in a limit of 165,000 pounds tractive effort at 15 mph. From 15 to 26 mph the limit is determined by the continuous rating of the main transformer. From 26 to 35 mph the speed is allowed to increase until motor current drops to 500 amperes. The next notch on the controller then causes shunting of the motor fields, following which the higher taps of the transformer bring speed up from 35 to 48 mph. Above that point the normal current versus speed curve of the motor is again followed.

The two *GG-1* units show the maximum hp of 12,150 at 39.5 mph (Fig. 2) on a short-time basis, and 10,000 hp at 53 mph on a continuous basis. The diesel, by contrast, provides a uniform pattern from 13 to 61 mph, reflecting the constant output of its prime movers. The hp characteristic of the rectifier locomotive represents a compromise between the other two. Waiting for traction motor current to drop before going into weak field operation appears to waste some of the locomotive power. However, in actual practice, very little time is lost because of other factors which influence train movement.

For a given set of conditions the coefficient of adhesion will be the tractive effort which can be used, up to the point of slipping, divided by the total weight on those wheels transmitting the torque. With dry rail a coefficient of 35 per cent can be depended upon for starting a train. As the speed is increased, the dependable limit is reduced until, at about 50 mph, it is in the order of 17½ per cent. On sanded wet rail, 25 per cent for starting is about the maximum coefficient to be expected. Train speed can be maintained at 50 mph on sanded wet rail provided tractive force does not exceed 12 per cent of its weight on drivers.

Table II. Results of Road Tests from Enola to Morrisville, Pa.

	A-C, Per Cent	Ignitron, Per Cent
Locomotive efficiency at nominal continuous hp	74.1	82.5
Locomotive efficiency at maximum continuous tractive effort	60.0	79.0
Locomotive power factor at nominal continuous hp	82.9	82.0
Locomotive power factor at maximum continuous tractive effort	63.4	82.0

It is apparent from the adhesion curve (Fig. 3) that, except in starting, the high short-time rating of the *GG-1* is useful only on dry rail. Continuous rating ranges from about 13 per cent in starting to slightly under 12 per cent, well below the sanded wet rail limits at 50 mph.

The diesel locomotive has a continuous tractive effort which is not high enough to cause it to slip on sanded wet rail. It can be seen that only the short-time rating at slow speed extends above the adhesive limit of sanded wet rail. Maximum continuous tractive effort is exerted at an adhesive limit of 19 per cent.

The flat speed-tractive effort characteristics of the rectifier locomotive with all motors in parallel permits it to maintain higher adhesion than is practicable on the a-c motive power with two or three motors in series.

The rectifier locomotive's maximum continuous tractive force is applied at an adhesive limit of 22 per cent. Between speeds of 4 and 24 mph some slipping might be expected on sanded wet rail, but in actual service, slipping is a rare occurrence. With presently assigned tonnage ratings, however, this locomotive has sufficient power to maintain a speed above 25 mph on all ruling grades.

From well-established formulas the railroad has calculated the friction resistance of the three types of locomotives and compared them with a typical 30-ton freight car. See Fig. 4. The resistance increases with speed in roughly a linear

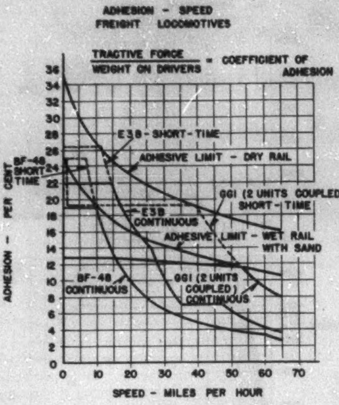


Fig. 3 (left). Speed rail-adhesion curves

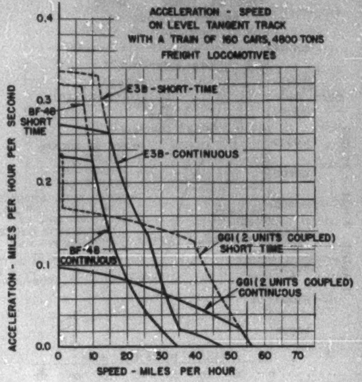


Fig. 5 (right). Speed acceleration curves

relationship of 160 pounds per 2.5 mph for all three locomotives. At starting speeds the difference between the rectifier locomotive, which has the lowest resistance of the three types, and the two *GG-1*'s is some 350 pounds, equivalent to about the resistance of two freight cars. In every case, to determine the net tractive force available for acceleration, the total resistance for locomotive and cars must be subtracted from total delivered tractive effort.

This net force applied to a train of known mass will produce a calculable acceleration. Fig. 5 shows what each of the three locomotives can do to accelerate a train of 160 cars (4,800 tons) on level tangent track. It is apparent that at speeds below 13 mph the rectifier, diesel, and *GG-1*'s rank in the order named. The ability of the *GG-1*'s to better use their hp at high speeds is shown by the fact that they overtake the rectifier at 24 mph on a short-time basis, and 31 mph on a continuous basis. The final balancing point for maximum speed is 35 mph for the diesel, 47.5 for the rectifier, and 56.5 for the *GG-1*'s. However, the maximum authorized speed for freight trains on the PRR is 50 mph.

Since actual operations call for much of the performance to be judged by ability on adverse grades, it should be noted that a locomotive hauling a train up a 1-per-cent grade will be required to expend 20 pounds of tractive force for each ton of weight of its train. From the relationship that force is equal to the product of mass and acceleration, it can be shown that this is equivalent to a negative acceleration of 0.2 mph per second. Thus, from the acceleration curves just examined it is apparent that at 10 mph the rectifier locomotive can haul the 160-car train up a 1.32 per cent grade, the diesel can haul it up a 1.13 per cent grade, and the *GG-1*'s up a short 0.81 per cent grade or a long 0.45 per cent grade. The *GG-1*'s or the diesel would be moving the train at 20 mph and the rectifier locomotive at 30 mph. Table II presents anticipated values from calculations and laboratory tests. They are compared with the expected values for a modern a-c locomotive at equivalent hp.

The real yardstick of performance which summarizes all features of the

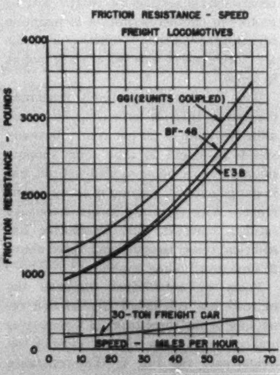
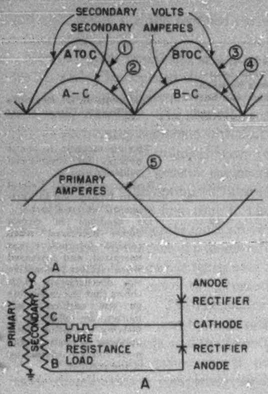


Fig. 4. Speed friction-resistance curves

Table III. Performance of Electric Freight Locomotives in Road Tests from Enola to Morrisville, Pa. (130 Miles)

	Class *GG-1*	Class *P5a*	Class *E2C*
Locomotive Data			
Type	a-c	a-c	a-c to d-c
Number of cabs	1	1	2
Continuous rail hp	4,620	3,750	6,000
Maximum speed, mph	90	70	63
Total weight, pounds	460,000	394,000	740,900
Weight on drivers, pounds	300,000	229,000	740,900
Over-all length, feet	79.5	62.7	124
Present tonnage ratings Enola to Morrisville:			
Adjusted tons (factor = 20)	6,000	6,300	16,800
Flat tons, 50-ton cars	4,280	4,500	12,000
Flat tons, 85-ton cars	4,850	5,100	13,600
Road Tests			
Date	8-23-46	8-14-46	2-19-52
Rail condition	dry	dry	dry
Number of cars	76	80	162
Adjusted tons	5,895	6,158	16,588
Flat tons	4,375	4,558	13,348
Per cent of rating	98.2	97.7	98.3
Time in motion	3 hours 28 minutes	4 hours 50 minutes	4 hours 20 minutes
Average speed, mph	37.4	26.8	30
Gross ton-miles	568,750	592,540	1,735,240
Gross ton-miles per train-running hour	163,905	122,679	400,440


SECONDARY VOLTS
SECONDARY AMPERES
A TO C
B TO C
①
②
③
④
A – C
B – C
PRIMARY AMPERES
⑤
A
PRIMARY
SECONDARY
C
B
PURE RESISTANCE LOAD
ANODE
RECTIFIER
CATHODE
RECTIFIER
ANODE
A

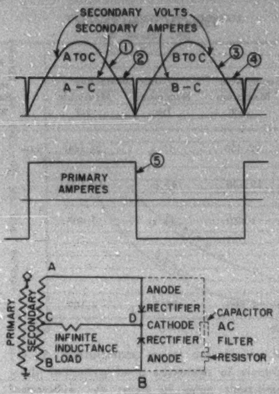

SECONDARY VOLTS
SECONDARY AMPERES
A TO C
B TO C
①
②
③
④
A – C
B – C
PRIMARY AMPERES
⑤
A
PRIMARY
SECONDARY
C
D
B
INFINITE INDUCTANCE LOAD
ANODE
RECTIFIER
CATHODE
RECTIFIER
ANODE
CAPACITOR
AC FILTER
RESISTOR
B

Fig. 6(A). Schematic illustrating effect of fully rectified single-phase power on pure resistance load

Secondary values 1 and 2 result from first half-cycle of primary amperes, 5. Values 3 and 4 result from second-half cycle of 5, etc.

Fig. 6(B). Schematic illustrating effect of replacing resistance load of Fig. 6A by infinite inductance load. Dotted lines indicate harmonic filter location

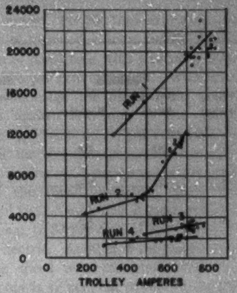


Fig. 8. Curves showing variation in I·T values with trolley amperes for each of four filtering arrangements

ability of any locomotive is the gross ton-miles per train-hour. As a final indication of this, Table III shows the results of road tests from Enola to Morrisville, Pa. (130 miles).

Since the data in Table III were taken, an even more convincing demonstration was made when the rectifier locomotive hauled 161 loaded coal cars over the same route at an average speed of 32 mph, giving the exceptional performance of 434,980 gross ton miles per train hour.

Telephone Circuit Influences

From the very beginning, the designers of rectifier-type motive power for railroads have recognized one potential handicap which could not be fully evaluated in advance. This is the inductive effect of harmonic frequencies in the audio range on communications circuits which operated in the general proximity of the trolley wire.

Experience with stationary ignitron substations had indicated fairly well-defined techniques and degrees of filtering to avoid objectionable telephone interference. Application of the ignitron to rolling stock, however, presented a new problem in the selection of adequate filtering equipment. It must not allow harmonic content to produce amplitudes at voice frequencies which would cause undue annoyance, nor could it be so large as to be impractical for mounting on a locomotive.

Some experience had been gained in the operation of the rectifier-equipped multiple-unit car prior to construction of the locomotives. With this as a guide an extrapolation to take into account the larger current ratings of the locomotives led to the decision to install on each locomotive unit a filter consisting of sets of 10 capacitors, of 23.5 microfarads (mfd) each (1,800-volt). Each set of 10 was connected in parallel across the anodes of the pair of ignitron tubes for a motor. Fig. 6(B) illustrates the schematic position of this a-c filter.

If a rectified sine wave of single-phase voltage is applied to a pure resistor load with no reactance in the voltage supply source, the current in the secondary will be half sine waves and the resultant line amperes will be a sine wave of current without harmonics. See Fig. 6(A).

If the pure resistor load is replaced with an infinite value of inductance as a load, the current wave will become rectangular and the line amperes will contain an infinite number of odd harmonics with a definite order of magnitude. See Fig. 6(B).

Thus it is seen that inductance in the d-c circuit produces harmonics.

The inductance in the d-c circuit generates a given magnitude of harmonics to produce a given wave shape of pulsating d-c amperes. These harmonic currents

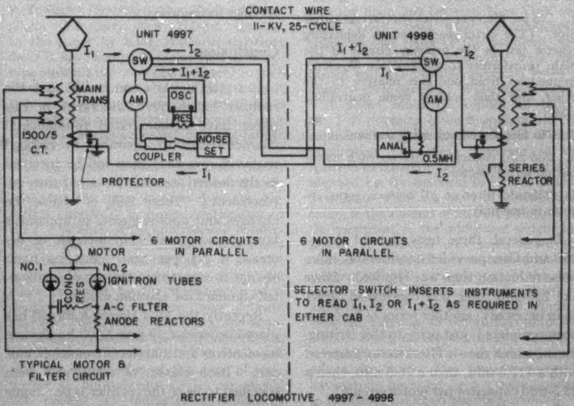

CONTACT WIRE
11-KV, 25-CYCLE
UNIT 4997
UNIT 4998
I_1
I_2
$I_1 + I_2$
SW
AM
OSC
RES
NOISE SET
COUPLER
MAIN TRANS
1500/5 C.T.
PROTECTOR
ANAL
0.5MH
SERIES REACTOR
6 MOTOR CIRCUITS IN PARALLEL
6 MOTOR CIRCUITS IN PARALLEL
MOTOR
NO.1
NO.2
IGNITRON TUBES
COND
RES
A-C FILTER
ANODE REACTORS
TYPICAL MOTOR & FILTER CIRCUIT
SELECTOR SWITCH INSERTS INSTRUMENTS TO READ I_1, I_2 OR $I_1 + I_2$ AS REQUIRED IN EITHER CAB
RECTIFIER LOCOMOTIVE 4997 - 4998

Fig. 7. Connections for noise measurements on rectifier locomotive

Table IV. Test Results

Test Number	Condition		Single-Unit Locomotive		2-Unit Locomotive		Average Db-Circuit *50304* while Locomotive was Between Safe Harbor and Bart		Comments on Db Readings in Phantom Circuit *50304*
	10-Cap Filter	Reactor	Maximum Db	Maximum I·T	Maximum Db	Maximum I·T	Metallic	Ground	
1	No	No	45.9	12,450	51.5	24,000	27	22	The db metallic is about average for commercial cable circuits.
2	No	Yes	45.5	12,150	45.5	12,150	24	16	Reduction in peak $I \cdot T$ about 2 to 1 reduction in average $I \cdot T$ about 4 to 1.
3	Yes	No	28.5	1,680	34.4	3,405	21	12	Db readings in circuit *50304* decreased when rectifier locomotive was energized, and increased when rectifier locomotive was de-energized. This shows that the capacitors on the rectifier locomotive are effective in filtering out other noise voltages as well as the noise voltages originating in the rectifier.
4	Yes	Yes	29.5	1,890	30.7	2,160	21	11	

flowing through the reactance of the transformer secondary winding produce a transformer volts per turn that causes harmonics to flow in the supply circuit.

If a low-impedance circuit is provided for these harmonics to flow without causing voltage drops in the transformer winding, it should be possible to reduce the harmonic content of the line current.

The anode-to-anode filter capacitor, Fig. 6(B) provides such a low impedance path for the harmonics to flow. The effect of this anode-to-anode filter can be considered in two ways, as follows:

1. The anode-to-anode filter short circuits the transformer secondary for the higher harmonics and thus prevents or reduces the harmonic volts per turn with a corresponding reduction in the harmonic content of the line current.
2. Assume an instant when voltage A is higher than C with some harmonic, say the 41st, flowing from A to D to C into the transformer winding at the mid-point of the secondary.

When the harmonic enters the transformer winding, it has two paths available for returning to A:

1. One from C through transformer winding to A.
2. The other from C through other half of transformer winding to B, then through a-c filter to A.

The impedance of the filter capacitor is very low at the 41st harmonic, so the harmonic current will tend to be equal in each path. Hence, the ampere turns produced by the harmonic current flowing from C to A will be cancelled by the ampere turns produced by the harmonic current flowing from C to B, thus leaving no harmonic ampere turns to reflect harmonics into the line current.

In the following is presented a brief analysis of the field methods which were used as an approach to determine the most practical value of filter, all things considered.

It should be pointed out that the commercial ramifications are many. Not only are private communications lines of the PRR involved, but the Bell Telephone System also has exposed sections in service at various places near the railroad right of way. This meant that several operating companies of the Bell System, as well as the Bell Telephone Laboratories and American Telephone and Telegraph Company, were all necessarily parties to the planning and execution of field tests.

In the first series of tests the 2-unit locomotive made four runs under the following conditions:

1. No a-c filters or series reactors on either unit. (The series reactor was a dephasing reactor put in the primary circuit of the main transformer to displace the harmonic frequencies from one cab relative to the other, and thus achieve some neutralization.)
2. No filters, one unit with reactor.
3. 235-mfd capacitor filter on each motor but no series reactor.
4. 235-mfd filter on all motors, series reactor in one unit.

In general, these tests showed that a 235-mfd filter provided considerably more noise reduction than was required. Since the 235-mfd filter was considered larger than necessary, another series of tests has been completed, just prior to this writing, during which the a-c filters were connected with a 117.5-mfd, then a 70.5-mfd, then a 23.5-mfd capacitor per motor circuit.

Two exhibits from the early tests only will show circuit arrangements during noise measurements as well as a few of the significant comparisons taken from observed data:

1. Rectifier locomotive no. *4997-4998*—connections for noise measurements. See Fig. 7. Note the noise measuring set (type *2-B*) and oscilloscope in one unit and harmonic analyzer in the other. By the use of current-selecting switches the test crew in either unit could pick up the transformer current from the other.
2. Current—telephone - influence - factor ($I \cdot T$) product versus trolley current. See Fig. 8. The almost linear relationship between total trolley current and $I \cdot T$ is shown here quite clearly.

The test results shown in Table IV cover noise measurements in a phantom circuit which probably has poorer than normal balance, but in spite of this only average commercial metallic decibel readings were obtained with the nonfiltered rectifier locomotive.

Conclusion

It is the writer's conclusion that the rectifier locomotives, described in this paper have demonstrated fundamental characteristics so desirable in railroad motive power that they will be specified as the desired standard for any future replacements. There will, of course, be changes and improvements in apparatus layout, and possibly an increase or decrease in hp per unit. But no basic change is expected in the use of ignitron-fed standard d-c traction motors.

Recently the New Haven Railroad has placed orders for 10 new 4,000-hp electric locomotives and 100 electric multiple-unit cars. Both classes of equipment were specified to be of the rectifier type. Some European countries, notably France, are also placing in service both locomotives and cars which use the rectified current for traction.

Certainly this practical proof of rectifier operation has opened the door to the use of commercial frequencies for railroad electrification. The resultant economy of such a step combined with that of using higher trolley voltages, now possible as a result of general engineering progress, points to an optimistic future for electrification for the first time in many years. Any such program must prove itself, of course, in the face of a dieselization which in some 15 years has taken over approximately 80 per cent of the locomotive hp in this country. The economics of investing capital on a unit-by-unit, rather than system-wide basis, and the flexibility of a self-contained prime mover have made dieselization a completely natural process. Even so, however, there exists in most cases a very close balance between the operating costs of a modern electrification and that of a similar dieselized operation—so close, in fact, that with sufficient traffic density and an adverse trend in the cost and availability of oil, one could easily see another large-scale trend to railroad electrification.

Not only is the cost of electric power less than that of fuel oil, for an equivalent amount of work, but the ability of an electric locomotive of any type to draw an unlimited trolley power offers a short-time maximum-hp service that is of extreme importance. For a practical example, consider the importance of speed in the PRR operation east of Harrisburg.

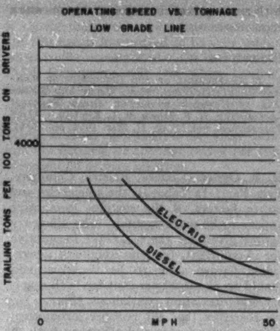


Fig. 9. Performance curves over Smithville grade indicating basic advantage of electric locomotive. A result of using short-time overload capacity

The operating "bottle-neck" on the east-west freight line is the so-called Smithville Grade, 22 miles in length. Operating speeds over this grade are an excellent criterion of the over-all locomotive performance from Harrisburg or Enola to the eastern terminals. Present class-*GG-1* and *-P5A* locomotives handle their assigned tonnage over this grade at over 30 mph.

Fig. 9 shows the relationship between trailing tons and speed over the Smithville grade. Per 100 tons on drivers, each type of locomotive may handle a maximum of 3,400 tons with a running factor of $16^{2}/_{3}$ per cent adhesion. With maximum tonnage the diesel speed is 12.5 mph, against 22.5 mph for the electric. This represents an 80 per cent improvement for the electric locomotive and a time saving of 55 minutes over the grade.

If the diesel operating speed were raised to 22.5 mph then it must reduce its tonnage to 53 per cent of that hauled by the electric. At 30 mph the present operating speed over the grade, the diesel tonnage is 47 per cent of that of the electric locomotive. That is, approximately 200 tons of diesel locomotive are required per 100 tons of electric locomotive to maintain the present schedule.

Clearly, there is much economy to be realized by the extension of railroad electrification in this country, and ignitron-rectifier motive power will be a major factor in making that economy a maximum.

References

1. The Pennsylvania Railroad Ignitron-Rectifier Locomotive, **C. C. Whittaker, W. M. Hutchison.** *AIEE Transactions*, vol. 71, pt. II, Jan. 1952, pp. 37–47.

2. Rectifier-Type Motive Power for Railroad Electrification, **L. J. Hibbard, C. C. Whittaker, E. W. Ames.** *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 519–24.

3. The Ignitron Locomotive, **C. C. Whittaker, W. M. Hutchison.** *Westinghouse Engineer*, Westinghouse Electric Corporation, East Pittsburgh, Pa., July 1952, pp. 122–26.

◆

Discussion

F. T. Garry (Southern New England Telephone Company, New Haven, Conn.): In the paper by King, Hibbard, and Gordon,[1] it has been pointed out that ignitron-type multiple unit cars and locomotives may develop $I\cdot T$ products of sufficient magnitude to cause objectionable levels of noise in paralleling communication circuits, if such units are not equipped with filters of proper design.

In the exposure areas that were chosen for test purposes, the data recorded indicate that the noise to ground in the disturbed circuits is in direct relationship to the $I\cdot T$ product of the ignitron unit in operation. The effect of various combinations of capacitors in reducing the $I\cdot T$ product developed has been shown, together with the effect produced by the introduction of a dephasing reactor. These tests indicate that with a filter, the $I\cdot T$ product, and hence the noise to ground in the telephone circuit, can be materially reduced. The conclusion that the introduction of this type of locomotive or multiple-unit (MU) car should not present an inductive co-ordination problem that cannot be satisfactorily resolved seems logical.

When The New York, New Haven and Hartford Railroad ordered 10 ignitron locomotives and 100 MU cars for use primarily on their main line between New York and New Haven, both their communication engineers and those of the Southern New England Telephone Company were concerned with the possible increased levels of noise on their circuits. While it was known that filters could be applied to reduce the $I\cdot T$ products of such equipment, two complications arose in the subject of filter application. The first, applying to the ignitron locomotive, is the weight limitation imposed by the entrance to Grand Central Terminal. Here we have a possible limitation to the amount of filtering strictly from the viewpoint of the additional weight imposed. A question also arises as to the cumulative effect of a combination of locomotives and MU cars operating in the same vicinity. The New Haven Railroad will be using many of the MU cars in tandem. How the influence of the various units will add, in effect, on communication circuits is probably a matter that will have to be resolved by tests. The present operating locomotives have not been a noise problem to the communication facilities involved.

In the New Haven Railroad area, we are not certain that the noise in the paralleling communication circuits will be at a satisfactory level after the tentatively scheduled filtering mentioned in the report. For example (granting that the exposures are not the same), although some of the metallic noise levels shown in the report would be satisfactory at the switchboard end of a trunk circuit, they would be objectionable at that level in the subscriber's receiver. We know from experience that metallic noise in a circuit that results in a receiver noise greater than 17 to 20 db may result in subscriber reaction.

The author suggests that consideration is being given to the use of this type of motive power on railroad lines where the a-c source of power could be from commercial 60-cycle transmission systems. Such usage could result in the harmonic frequencies generated in the railroad system feeding back into the power-distribution systems if adequate filters were not applied. This problem is one of vital concern to telephone companies due to existing exposure to such commercial power circuits and to the constant increase in the joint use of pole line plant by telephone and power company facilities.

REFERENCE

1. Rectifier Motive Power—Inductive Co-ordination Considerations, **E. B. King, K. H.**

Gordon, L. J. Hibbard. *AIEE Transactions*, vol. 73, pt. II, July 1954, pp. 107–18.

H. F. Brown (Consulting Engineer, New Haven, Conn.): This paper is a most timely addition to the existing published data on ignitron locomotive performance. The author is to be complimented on the clear and concise manner in which not only the performance but the design data have been presented.

An important feature of the design, that of having all the motors in parallel at all times, is stated in one single sentence. This feature is worthy of emphasis, for it gives this locomotive very desirable and unusual operating characteristics worthy of adoption in all future locomotives of the rectifier type.

The table showing the comparative efficiency and power factor of this ignitron-type locomotive, compared with these same characteristics for a-c commutator motor type is quite important as there have been conflicting ideas on both these factors not only in this country but abroad. Are these power-factor values with or without the harmonic filtering apparatus? If without, what effect on the power factor of the locomotive do the various amounts of filtering equipment described in the various tests have?

A very important consideration is the necessity of designing the circuits and the harmonic filtering apparatus so that the operating power factor of the locomotive is not seriously impaired. This should be kept in mind not only by the engineers having the traction performance of this type of locomotive in view, but by the communication engineers as well, whose main interest may be that of obtaining the very minimum amount of inductive disturbance. There must always be an economic compromise by both interests in this matter.

In Fig. 8 are shown the various $I\cdot T$ values for various trolley amperes. With a higher voltage, say a nominal voltage of 25 kv, the same train loads could be moved with half the number of amperes. If it were possible from tests made to estimate or calculate the effect of a 60-cycle fundamental frequency, instead of 25 cycles, on the telephone-influence factor, the $I\cdot T$ values for commercial frequency with the lower current values might be much lower with 50 or 60 cycles than those shown for 25 cycles.

In Table III, it would have been interesting to compare the "gross ton-miles per train running hour" of the rectifier locomotive with the performance of diesel units having the same number of traction motors, and with the same maximum trailing tons. The speed, and therefore the "ton-miles per hour" would be appreciably less.

In the *Conclusion*, third paragraph, it is stated that "there exists in most cases a very close balance between the operating costs of a modern electrification and that of a similar dieselized operation." It is beginning to appear that this balance is not as close as has been thought. Factors not mentioned in, or within the scope of this paper, but which have been mentioned in another paper[1] are beginning to show electrification in a much more favorable light. In the last two sentences in this paper, the author very concisely sums up the comparison between diesel and electric operation.

It is certain that this paper will be studied with great interest by railway electrification engineers abroad.

REFERENCE

1. A REAPPRAISAL OF THE ECONOMICS OF RAILWAY ELECTRIFICATION: HOW, WHEN, AND WHERE CAN IT COMPETE WITH THE DIESEL-ELECTRIC LOCOMOTIVE, H. F. Brown, R. L. Kimball. *AIEE Transactions*, vol. 73, pt. II, Mar. 1954, pp. 35–51.

H. S. Ogden (General Electric Company, Erie, Pa.): In Table II, the efficiencies and power factors of the rectifier locomotive appear to be quite factual. The efficiency figures given for the a-c *GG-1* locomotive are about right, but corresponding power-factor figures seem to be in error, and should be about 90 per cent for both nominal and maximum continuous effort. We see therefore that the rectifier locomotive, with a power factor and efficiency both in the neighborhood of 82 per cent, may have a somewhat better efficiency and a somewhat worse power factor than the corresponding straight a-c locomotive.

The paper describes the use of a "dephasing reactor" on one locomotive unit. Referring to Table IV the benefits of the use of this reactor as compared to the capacitor filter are summarized. Where no capacitor filter is used, the reactor is of some benefit. Where the capacitor filter is used, the dephasing reactor gives very little further improvement. The basic improvement due to the reactor, as indicated by measurements of decibels on the telephone circuits, is not nearly as great as the improvement obtained by the use of the filter capacitor. In this connection, I note that there is a difference in the decibel readings given in this paper when compared to reference 1 of Mr. Garry's discussion for the same conditions of operation. The data involved are given in Table I of the latter paper, *First Series*, and the differences are considerable. I do not believe, however, that these differences would alter the conclusion one would draw from the two sets of data taken individually.

If, for a given rectifier locomotive you have a typical primary a-c wave with its harmonic analysis, it can be shown analytically that the improvement to be expected from the dephasing reactor would not equal the improvement obtained by the use of the resistor-capacitor circuit. I might go further, and say that even if arrangements were made to dephase each of the six motor circuits on one locomotive unit, the resulting improvement in the telephone-influence factor would probably not be worth the effort, and would not be as good as that achieved by the resistance-capacitor circuit connected across the rectifier anodes.

The author's remarks concern a locomotive in which a diametrically connected rectifier is used. A bridge-connected rectifier will behave in exactly the same manner.

D. R. MacLeod (General Electric Company, Erie, Pa.): The first part of this paper appears to be an argument in favor of a freight locomotive having this horsepower per ton of locomotive weight. By carefully choosing the ground for comparison, he shows that a high-horsepower *GG-1* locomotive produced approximately the same gross ton-miles per train running hour per ton of weight on drivers as the low-horsepower *E2C* locomotive geared for a maximum speed of 65 miles per hour. The *GG-1* locomotive used in the test is a passenger locomotive and, therefore, is restricted when operating in freight service. If it had been built for freight service and geared for the same maximum speed and weight on each driver as the *E2C*, it would have had a tonnage rating of 8,000 adjusted tons, and would be able to develop 90,000 pounds tractive effort up to 38 miles per hour. Two of these locomotives would have produced approximately the same ton-miles as the *E2C* over the section of track on which the tests were made, and the ton-miles per train running hour would have been 40-per-cent greater. They could produce even more ton-miles per train hour than the *E2C*, were it not for the speed restrictions that govern trains on the Atglen and Susquehanna Branch, where the tests were made. On the main lines of the electrified divisions, where higher speeds are permissible (except for mineral trains), the higher horsepower locomotive would have shown a marked advantage over the lower horsepower locomotive with approximately the same weight on drivers.

In this paper, the tonnage ratings and performance comparisons of the several locomotives are considered on the basis of operation over a 130-mile section of line including the 0.3-per-cent ruling grade between Columbia and Parkersburg. The comparison is particularly unfavorable to the *P-5A* locomotive and, in our opinion, quite misleading. A *P-5A* locomotive with 225,000 pounds on its drivers is rated 4,350 flat tons over this grade. It can move its rated tonnage up the ruling grade at 40 miles per hour, whereas the *E2C* can move its tonnage up the grade at only 20 miles per hour, see Fig. 1. On the New York Division, where the speed restriction for other than mineral freight is 50 miles per hour, the *P-5A* locomotive can move 4,350 tons of 60-ton cars on level tangent track at 50 miles per hour. The *E2C* cannot move its rated tonnage at this speed on the level. If a test had been made with three *P-5A* locomotives and the 162-car 13,348-ton train with which the *E2C* was tested, it would have been found that the scheduled speed on the New York Division would have been much higher than when the *E2C* locomotive was used. Using Davis friction for its 82-ton cars, it may be shown that the *E2C* would require approximately twice as much horsepower as it has available at 50 miles per hour, as shown in Fig. 2. Three *P-5A* locomotives could haul the same train at 50 miles per hour on the level without exceeding the *AIEE* temperature limits for continuous operation.

Contrary to the impression given in the paper, the *P-5A* uses its high horsepower to good advantage on the Pennsylvania Railroad. This is indicated by the fact that the PRR gets approximately twice as many miles per month per locomotive out of its fleet of *P-5A* locomotives as it gets out of the *E2C*. Part of this is because the *E2C* type of locomotive is best suited for hauling heavy mineral trains that are restricted to relatively low speeds under existing operating rules.

The question is often asked, how much horsepower per ton of weight on drivers can be used to advantage in freight service? The wet-rail adhesion curve shown in Fig.

Table V. Usable Versus Available Horsepower

Miles per Hour	Usable Horsepower Per Ton on Drivers For Worst Rail Conditions	Available Horsepower Per Ton on Drivers			
		E2C	P-5A	E2B	Diesel BF-48
10	10.1	14.9	16.4	11.4	9.9
15	14.0	17.3	23.2	17.1	10.7
20	17.0	18.7	31.1	23.0	10.6
30	22.7	16.8	46.3	34.3	10.4
40	27.6	15.2	49.5	33.8	10.4
50	32.0	17.8	40	27.8	10.4

3 can be used to answer this question. This is a good curve to use, except under some very bad conditions that the writer has experienced in the West. It may be translated into usable horsepower at the rail per ton on drivers under the worst rail conditions for any axle loading.

Table V shows that a freight locomotive that is to enable a 50-mile-per-hour railroad to use its tracks to the best advantage under all weather conditions should have 32 horsepower per ton of weight on drivers (38-engine horsepower for diesel-electrics). The *P-5A* locomotive is the only one of the above locomotives that fully meets this requirement. It has more horsepower than is usable under the worst rail conditions, but this excess has proved to be very useful when traffic is congested. On bad days, a *P-5A* locomotive will slip when hauling tonnage trains unless the speed is lowered to the point where it can hold the rail. On these days, it may have to go over ruling grades at lower speeds than the other locomotives because its tonnage rating per ton of weight on drivers is higher.

The thesis that the economical movement of freight under conditions of high traffic density does not require high horsepower per ton of weight on drivers will destroy the chief argument justifying the admittedly high cost of electrification, if allowed to go unchallenged. We believe that electric locomotives with low horsepower per ton of weight on drivers cannot compete with the diesel-electric locomotive on a busy railroad. An electric locomotive capable of developing at least 30 horsepower per ton of weight on drivers is required, because it will take three times the weight of diesel-electric locomotive to match the horsepower performance of the electric. On a horsepower basis, the electric locomotive has a 2-to-1 advantage in maintenance cost over the diesel-electric, and it is chiefly on this basis that the straight electric is outstanding.

F. D. Brown: The constructive comments received from all discussers are sincerely appreciated. Those applying to the performance comparisons of several locomotive types were particularly appropriate. In this respect, however, as pointed out in the paper, my intent was to present a report of the operating results of existing locomotive types on an existing railroad system. Unquestionably, even better performances could be obtained by the design of electric locomotives to meet certain sets of conditions.

The power-factor and efficiency figures given for the a-c locomotives are calculated values based on designs not actually in existence. The values presented for the rectifier locomotives are accurate within a reasonable tolerance. Qualitatively, it can be said that for a given load, the efficiency goes down, due to the losses in the damping resistors as the degree of harmonic filtering is increased. On the other hand, the power factor is improved because of the effect of the added capacitors. I regret that we cannot indicate at this time accurate values for these trends with respect to the degree of filtering. However, a considerable amount of work is continuing on this subject and there will be data published on it in the future.

Series- Versus Parallel-Connected Generators for Multiple-Engine D-C Diesel-Electric Ship-Propulsion Systems

JAMES A. WASMUND
MEMBER AIEE

THE USE of electric equipment in a Diesel propulsion system for ship drive serves as a nonmechanical means of connecting one or more relatively high-speed engines to an inherently slow-speed propeller. Such an installation is much more flexible than a mechanical connection since the prime movers can be physically distributed to best utilize the available space and can also be instantly connected to or disconnected from the propulsion circuit. The use of synchronous a-c machinery has a number of attractive features. The motors and generators are smaller and lighter, have higher efficiencies, lower maintenance, and lower first cost than equivalent d-c machines. However, there are also some disadvantages. The speed range of the propelling motor is limited to that of the prime mover because there is a fixed ratio between generator and motor speed. Also, there is no means of varying the motor speed-torque ratio such as can be obtained by using d-c machinery. Thus, the use of a-c Diesel-electric drive is limited to vessels making relatively long voyages at a more or less fixed speed and where maneuvering is of relatively small importance. Direct current, on the other hand, has almost ideal characteristics for use in supplying propelling power to tugboats, ferries, dredges, and ice-breaking vessels, which are almost continuously maneuvering and operating under varied conditions of loading.

This paper does not cover a wide field of application; the problem discussed is limited to very few vessels, at least at the present time, since most Diesel-electric propulsion systems are limited to two or, at most, three prime movers per screw. However, the problem is not purely academic since the example chosen in making the calculations is an actual installation.

Of the many multiple-engine d-c Diesel-electric-propelled ships now in operation, almost all use the series loop system wherein the generator armatures are connected in series; likewise, if there are multiple motors driving the propeller in tandem or through a gear unit, the motor armatures are also in series in the loop circuit and are interspersed between generator armatures in order to keep the ground potential to a minimum. In most cases the selection of the series connection is the proper one and can be justified by a simple analysis. However, there is a strong tendency to propose the series loop system without careful study. If such a study were made, it would in some cases lead to the conclusion that the generator armatures should be connected in parallel.

The series loop circuit has many desirable features. One of the most important is that, up to a limiting value of motor field weakening, it is possible to utilize the full rated horsepower of all available engines regardless of the number of engines in operation; another is that load division between the various engines is satisfactory regardless of minor differences in engine speeds or generator

---

Paper 54-144, recommended by the AIEE Marine Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 16, 1953; made available for printing December 14, 1953.

JAMES A. WASMUND is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

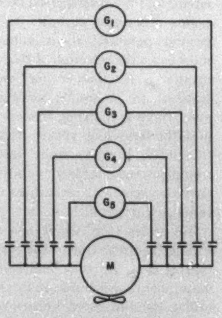
Table I. Calculations to Determine Available Shaft Horsepower Under Full Running and Stalled Ship Conditions with Different Combinations of Engine Generator Sets in Use

Based on an Arrangement of Five Generators and One Motor, a Constant Horsepower Motor Speed Ratio of 1.46 to 1, and a Motor Field Weakening Limit of 40-Per-Cent Flux

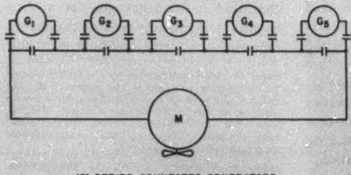
%	% BASE RPM			% GENERATOR FIELD FLUX			% GENERATOR VOLTS			% GENERATOR CURRENT			% MOTOR FIELD FLUX			% MOTOR CURRENT			% PROPELLER TORQUE			% MOTOR HORSEPOWER		
E	Pa	Pb	S	Pa	Pb	S	Pa	Pb	S	Pa	Pb	S	Pa	Pb	S	Pa	Pb	S	Pa	Pb	S	Pa	Pb	S
FREE RUNNING																											
100	146	146	146	100	100	100	100	100	100	100	100	100	68.5	68.5	68.5	100	100	100	68.5	68.5	68.5	100	100	100
80	134.5	134.5	134.5	100	100	100	100	100	100	100	100	100	74.3	74.3	59.5	80	80	100	59.5	59.5	59.5	80	80	80
60	123	123	123	100	100	100	100	100	100	100	100	100	81.4	81.4	48.8	60	60	100	48.8	48.8	48.8	60	60	60
40	107.5	107.5	100	100	100	100	100	100	100	100	100	100	93	93	40	40	40	81.3	37.2	37.2	32.5	40	40	32.5
20	85.5	79	50	85.5	100	100	85.5	79	100	117	100	100	100	100	40	23.4	20	20	23.4	20	8	20	15.8	4
STALLED SHIP																											
100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
80	93	89.5	93	93	100	100	93	89.5	100	107.5	100	100	100	100	86	86	80	100	86	80	86	80	71.6	80
60	84.3	76.5	84.3	84.3	100	100	84.3	76.5	100	118.6	100	100	100	100	71	71	58.5	100	71	58.5	71	60	44.7	60
40	73.7	62	73.7	73.7	100	100	73.7	62	100	135.5	100	100	100	100	54.3	54.3	38.4	100	54.3	38.4	54.3	40	23.8	40
20	58.5	41	50	58.5	100	100	58.5	41	100	171	100	100	100	100	40	34.2	16.8	62.5	34.2	16.8	25	20	6.9	12.5

E—Number of engine generator sets in operation.
Pa—Parallel-connected generators where engine overload is prevented by decrease in generator field flux.
Pb—Parallel-connected generators where engine overload is prevented by decrease in engine speed.
S—Series-connected generators.

voltages. With the series system, the only effect of such differences is a proportional difference in the load carried by the generators and their respective engines, whereas parallel operation of generators with the variable voltage system requires much more precise automatic engine governing, matching of generator load characteristics, and manual vernier adjustment of engine speed or generator voltage to insure equal load division.



(A) PARALLEL CONNECTED GENERATORS



(B) SERIES CONNECTED GENERATORS

Fig. 1. Schematic setup switch arrangement for series- and parallel-connected generators for an installation utilizing five propulsion generators and one propulsion motor

Series-Connected Generators

With generators connected in series, as shown in Fig. 1(B), a reduction in the number of engines in use brings with it a proportional reduction in voltage applied to the motor terminals and this in turn, if uncompensated for, results in a proportional reduction in motor speed. Since the power required to drive most propellers varies very nearly as the cube of the speed, it is desirable to weaken the motor field to such an extent that the motor will return to a speed which will utilize the full available engine horsepower. Field weakening of the motor is effective over a wide range; however, reduction to less than about 40-per-cent field flux is usually not permissible since it can result in bad commutation and poor motor performance. In the example chosen it will be seen that the minimum flux condition will actually limit the available shaft horsepower when less than 50 per cent of the engines are in use.

Parallel-Connected Generators

In contrast with the series case, for generators connected in parallel, as shown in Fig. 1(A), a reduction in the number of engines in use does not reduce the voltage applied to the motor terminals and, if uncompensated for, the motor will attempt to maintain full speed. This of course will result in overloading of the connected generators and engines. Therefore it is necessary to increase motor field flux sufficiently to slow down the motor and thus match required propeller horsepower with available engine horsepower or, if the motor field cannot be further increased, it becomes necessary to decrease the generator voltage.

Means of Decreasing Generator Voltage

Generator voltage can be decreased in one of two ways. Engine speed can be reduced while maintaining full generator field flux. However, this results in less available engine horsepower output and, therefore, less shaft horsepower output than a direct proportion to the number of engines in operation. Whether or not this decrease in shaft horsepower is of prime importance is a debatable question. In any case, the reduction can be readily calculated and is shown in Table I and in the figures. Generator voltage can also be decreased, while maintaining full engine speed, by reducing the generator field flux. This method has the advantage of utilizing the full engine horsepower available. However, to take advantage of the full engine rating and to main-

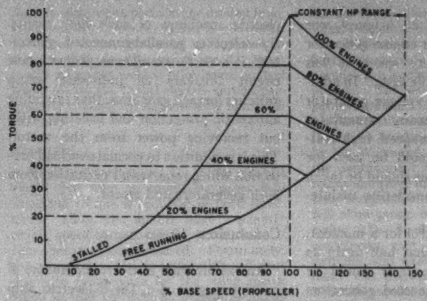


Fig. 2. Propeller speed-torque curves from free running to stalled ship conditions based on an arrangement of five generators and one motor; a constant horsepower motor speed ratio of 1.46 to 1; and a motor field weakening limit of 40-per-cent flux

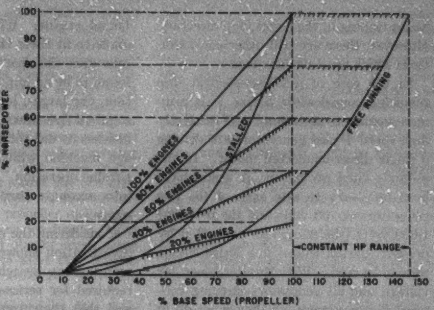


Fig. 3. Propeller speed-horsepower curves from free running to stalled ship conditions based on an arrangement of five generators and one motor; a constant horsepower motor speed ratio of 1.46 to 1; and a motor field weakening limit of 40-per-cent flux

tain full generator output, it is necessary for the generator to carry a current overload to compensate for the reduction in voltage. If this current overload is not carried to extremes and is only for short-time operation, the generator can carry the increased current at a moderately increased temperature rise without damage. By referring to Table I for the example chosen, it can be seen that, with 80 per cent of the engines in operation, a 7.5-per-cent increase in armature current will permit full 80-per-cent motor output as compared to 71.6-per-cent motor output by reducing engine speed. It should be noted that this does not result in overcurrent for the motor. Actually, the motor current is only 86 per cent in this particular case. As more generators are taken out of service, the current overload on those remaining becomes greater and a point will be reached where full remaining engine horsepower cannot be utilized unless the generators are originally designed with excess capacity.

Reasons for Selecting D-C Drive

In most cases where direct current is selected for ship drive, one of the reasons for the selection is its ability to provide increased torque at low speed as illustrated in Fig. 2, in other words, constant horsepower over a speed range as illustrated in Fig. 3. This is especially true for vessels which are used for tug service, towing, or ice breaking. For this reason, most d-c ship propulsion motors are designed to have a considerable range over which the speed can be varied by field control. This is an important point to consider when deciding on series- versus parallel-connected generators. For purposes of comparison, an example has been chosen which represents an installation consisting of five generators driving a propulsion motor. The weak field motor speed is 46 per cent greater than the full field or base speed.

Open-Water Running

If a vessel with parallel-connected generators is running in open water without a tow, it is possible to reduce the motor speed by field strengthening and to utilize full engine output with as few as two of the engines in operation without any reduction in generator voltage. With only one of the engines in operation and generator voltage reduced to 85.5 per cent, full output of the engines can be utilized by carrying a 17-per-cent generator overload current. If the vessel were to have series-connected generators with the motor flux necessarily limited to a minimum of 40 per cent, then only 81.3 per cent of the connected engine rating can be utilized with two of the engines in operation; with one engine in operation, only 20 per cent of its rating can be utilized. Thus it can be seen that, for free running, the series-connected plant operates at a disadvantage when under 1- and 2-engine operation. Actually this limitation could be considered academic and of no practical importance.

Stalled Ship Operation

Under maximum towing or ice-breaking conditions when the ship is completely stalled, the picture is somewhat different. The series system can go from five to two engines in operation without sacrifice in the utilization of full engine horsepower. With one engine, only 62.5 per cent of the available power can be absorbed by the motor because of the limitation in field weakening. Under the same conditions with parallel-connected generators and generator field weakening, full available engine power can be utilized with four engines and 7.5-per-cent generator overload current, three engines and 18.6-per-cent generator overload current, two engines and 35.5-per-cent generator overload current, and one engine and 71-per-cent generator overload current. Overload current up to 20 per cent can usually be tolerated for short periods of time without materially shortening insulation life provided the overloads are applied rather rarely. Using three engines at their maximum rating will not, in general, result in oversize generators. However, operation with one or two engines at their maximum rating will definitely require oversize generators.

If the generator field flux is maintained and voltage reduced by reducing engine speed, the available output drops from a possible 80 per cent to 71.6 per cent with four engines, from 60 per cent to 44.7 per cent with three engines, from 40 per cent to 23.8 per cent with two engines, and from 20 per cent to 6.9 per cent with one engine. It should be remembered that these figures represent a stalled ship condition, which is the worst possible condition encountered, for as soon as the hull begins to move through the water the load begins to decrease for the same propeller speed.

Additional Considerations in Selecting Generator Connection

From the foregoing considerations it is evident that each system has inherent advantages and disadvantages with the scales tipping in the series loop direction. However, there are still other considerations which must be examined.

In the case of a Diesel-electric ship requiring considerable power, it is common practice to use a number of relatively high-speed Diesel generator sets driving a single direct-connected motor. In the example cited previously, it would have been necessary to rate each of the five generators at 200 volts and the motor at 1,000 volts (the maximum permitted by present Marine Rules) if a series circuit were to be used. With this low generator voltage and all generators carrying full-rated motor current, as they must do in a series circuit, the generator commutators would have become so large that it would have been necessary to go to a larger diameter generator frame than otherwise required. Because of centrifugal forces and design limitations in commutator speed, the larger diameter armature could not be run at the rated speed of a modern high-speed Diesel engine. This made it necessary to consider the use of a relatively low speed and correspondingly larger and heavier Diesel engine. When an attempt was made to fit these larger engine generator sets into the machinery space, it was found to be an impossible task. In addition, the larger, lower voltage generator required five times as many commutator brushes as did the high-speed, high-voltage machine which would be used for parallel operation. This should be taken into account when considering maintenance.

Considering the control for a moment, investigation showed that half again as many setup switch contacts would be required for series-connected generators and that they would have to have five times the current rating of those required for parallel-connected generators. The particular installation being studied was for approximately 10,000 horsepower. In low power installations, this would not be of major importance. However, in this particular case, it would have required special development of high current contactors because of the high current rating.

The potential energy which it is possible to feed into a fault also becomes a serious consideration in an installation involving electric machines of such high rating. By selecting parallel-connected generators, it is easily possible to utilize 2-pole circuit breakers as individual setup switches for each generator, thus providing automatic protection and isolation without removing power from the motor. This is in addition to normal overload protection which removes all excitation from both generators and motor.

Conclusion

An examination of the requirements of multiple-engine d-c Diesel-electric ship drive will, in most cases, result in the selection of the series loop circuit. However, there are special cases where parallel-connected generators will offer a sufficient number of advantages to dictate that this circuit be used. In most cases the answer will be obvious without too much study, but for vessels requiring more than 2,000 horsepower in a single armature motor and utilizing three or more engines, a study should be made in each case to make certain that the system offering the most advantages is installed.

◆

Discussion

W. H. Fifer (Bureau of Ships, U. S. Navy Department, Washington, D. C.): Mr. Wasmund's paper presents a clear picture of the basic characteristics of parallel-connected generators versus series-connected generators for d-c Diesel-electric ship-propulsion systems. However, the basic reason for employing parallel-connected generators on the ice breaker currently under construction was the inability to provide the same number of propulsion motor armatures as propulsion generator armatures employed. Due to lack of space and the questionable ability of reduction gears to stand up in ice breaker service, a single-armature direct-connected d-c propulsion motor was selected. This decision resulted in parallel connection of generators for the reasons outlined by Mr. Wasmund. Some reduction in fault protection results from parallel connection of propulsion generators, since a fault on the generator side of a propulsion-generator circuit breaker would require the circuit breaker to interrupt the short-circuit current of all the other propulsion generators connected plus the short-circuit current provided by the propulsion motor. This current would probably be well in excess of the breaker rating, which would destroy the breaker and might cause the other connected machines to flash over. This calculated risk is minimized by locating each circuit breaker close to its respective generator.

For ice breaker service the maximum propeller thrust is of primary importance and in the event one or more generators are not available, it is desired to have the full available horsepower from the remaining engines. It is believed the best balance for accomplishing this condition is to overbuild the generators so that they may continuously carry some overcurrent and to overbuild the motor by increasing the magnetic iron. The balance between the additional current-carrying capacity of the generators and the additional motor iron should be such that full engine output can be continuously obtained from three engines when operating under ice-breaking conditions.

H. C. Coleman (Westinghouse Electric Corporation, East Pittsburgh, Pa.): This paper gives a clear and concise comparison of the operating advantages and disadvantages of series- and parallel-connected d-c generators. However, it would appear that, for the example cited, machine and control equipment design limitations and expediency, as well as ship space restrictions, played the dominant part in leading to the selection of the parallel system. This will usually be the case where so much power is involved. In every case the operating comparison must be worked out to make certain that the parallel system will permit reasonable, reliable operation at all ship design conditions. It would be a most unusual and rare case where the parallel system would not be entirely satisfactory.

From the ship design standpoint, a very important advantage of the parallel system is the considerable saving in size, cost, and weight of main cable. In the case cited in the paper, the main generator circuit current to be handled is less than one-fourth that for the series loop system.

In the early days of Diesel-electric ship-propulsion application, there was a great deal of discussion of possible parallel operation of generators but on final analysis the series system was almost invariably adopted. A major contributing factor was the poor governing characteristics of the engines of those days. A few attempts were made to parallel the auxiliary generators which were direct-connected to the propulsion generators but this proved impractical. Great improvements have been made in governor design and the modern Diesel engine has excellent characteristics.

Except for ships with special requirements, such as ice breakers, cable layers, large dredges, and ferry boats, it is believed that vessels requiring Diesel-electric propulsion plants of 5,000 shaft horsepower and above will use a-c machinery. It will give entirely satisfactory service and has considerable advantage in the way of lower first cost, lower weight, smaller size, higher efficiency, greater simplicity and ruggedness, and lower maintenance. With that system we automatically have parallel operation.

W. Schaelchlin (Westinghouse Electric Corporation, Buffalo, N. Y.): Mr. Wasmund has presented some very interesting facts about the two systems of Diesel-electric propulsion. Each one appears to have its merits and it should not be difficult to select the most suitable arrangement for a given application.

I would like to add some information about the operating characteristic of the two systems, such as reliability, stability, engine protection, and other features of interest to the operator.

RELIABILITY

Both systems are eminently successful in this respect. A minimum of control equipment is required to perform the desired functions and the propeller speed is controlled by adjusting the generator voltage and in certain cases also the engine speed. Practically all drives include pilot-house control which makes the Diesel-electric system particularly suitable for maneuvering in narrow channels.

STABILITY

All propulsion motors are of the straight shunt type with no stabilizing series field since full torque is required in both directions. Thus, the motors have a slightly rising speed-torque curve, particularly if they are not compensated. Fortunately, the generator can be designed with sufficient droop to compensate for the rising motor speed so that the over-all system is stable. This comment applies equally to both series- and parallel-connected generators.

LOAD BALANCE OF PARALLEL GENERATORS

There is no particular problem in operating generators in parallel. By providing them with a droop of about 8 per cent, they will share the load very satisfactorily. The generator designer has to make sure, however, that there is also positive droop at the low-voltage points. In certain cases, it may happen that the overcompensation of the commutating poles may actually cause a slightly rising voltage characteristic which, of course, is not satisfactory for parallel operation.

The performance of the engine governor is also of importance. It is relatively easy to provide from 3 to 5-per-cent droop to insure additional load balance, if required.

ENGINE PROTECTION

There is the danger of engine reversal in case of loss of engine torque with the generators connected in series. This is particularly true when there are more than two engines. Thus, it is considered desirable in certain cases to add special engine failure protection.

The parallel-connected generators do not have this limitation. On the other hand, it is possible to overspeed an engine by simply weakening the generator field. This will cause the generator to motorize and possibly result in overspeed. Thus it is necessary to provide overspeed protection.

SWITCHING

Both series- and parallel-connected generators require relatively simple switching. However, the parallel-generator scheme is more flexible, particularly for large installations. Moreover, it is possible to switch under load and add generators to the bus without momentary loss of propulsion.

REGENERATION DURING MANEUVERING

Very often the question arises as to possible overspeeding of the engine-generator during maneuvering, i.e., when regenerating during slowing down of the propeller. Experience has proved that there need be no concern. In the first place, the stored energy of the propulsion motor and propeller is relatively small compared to that of the higher speed generators. Moreover, there is a positive propeller load torque that helps to slow down the motor and therefore absorbs some of the stored energy.

H. H. Curry (Commander United States Coast Guard, Smyrna, Ga.) (Retired): Mr. Wasmund's much needed paper has fairly compared features of the series versus parallel d-c systems of Diesel-electric propulsion. The setup of three generators in parallel on each main motor was selected for the first class of 10,000-horsepower Coast Guard ice breakers for the reasons outlined by Mr. Wasmund, primarily as best meeting the requirement that, with three engine rooms and three motor rooms, any two adjacent compartments might be flooded and still allow power on a propeller without entering the flooded compartments. The complexity, cost, and size of series setup switches and cables to meet this requirement was excessive, although easily met with the parallel system.

A feature not mentioned by Mr. Wasmund is that if an engine in a series circuit looses power it cannot unload itself by slowing down. If in a 5-engine series installation one engine cuts out, its armature current remains high and in the same direction, and if excitation is not changed generator torque may run the engine backward. Should an engine in a parallel system lose power generator armature current will reverse and it will operate as a motor with the same direction of rotation with less chance of damage to the engine and its attached pumps. This may be important with a wet muffler or underwater exhaust.

The case selected for comparison indicates a rather large propeller for the power and speed. For the usual ice breaker or tug boat installation propellers are too often limited on diameter by draft or for protection and the change in motor revolutions between running free and stalled is less than the 46 per cent of stalled speed indicated. For instance, a propeller with pitch ratio of 1 might at 25-per-cent slip turn at 127 per cent of the 100-per-cent slip revolutions on the same power.

I cannot agree that governing is much more of a problem for the parallel than series system. Both have been demonstrated to be entirely practical for bridge control without requiring manual vernier adjustments with good commercial governors. I am of the opinion that the fuel limiting governor with a definite stop adjusted as a function of control position or engine speed to give the optimum maximum brake mean effective pressure engine speed curve is the preferable control system for the upper part of the power range. The parallel system is definitely better suited to operation with the fuel limited in this way, as an engine by slowing slightly can adjust its available torque to the torque required by the generator, while in the series setup a much greater change in speed is required to reduce engine load as, obviously, with all generator armature currents equal and fields equal the torque of an individual engine can be reduced only in proportion to change in main motor torque or approximately proportional to the two-thirds power of the total power supplied that motor. Stalling and reversed rotation of engines in series installations has been experienced with fuel limits and engines in poor adjustment.

With as many as five engines per motor I would closely study the advantages of alternating current. If operation at 100-per-cent slip were of sufficient importance I would even consider the use of more power to allow the same propeller speed at 100-per-cent slip with alternating current as with direct current. The saving in electric machinery cost and weight would probably buy the additional engine. Then an engine could be cut out when running free. At reduced power only the number of engines required for the desired propeller torque would be used so over-all efficiency of engines and transmission could be good over much of the power range. With alternating current, the power available for 1-, 2-, 3-, and 4-engine operation would only be about 6, 19, 42, and 72 per cent of 5-engine operation if allowance is made for operating at reduced BMEP at lower engine speeds. For frequent maneuvering direct current is preferable.

O. A. Wilde (Sun Shipbuilding and Dry Dock Company, Chester, Pa.): The first part of the paper concisely and accurately describes the basic considerations inherent in Diesel-electric propulsion applications. The paper also invites consideration of a particular set of conditions which make it desirable to review the choice of series- versus parallel-connected generators. I agree with the author's approach to this subject but I desire to comment upon certain points that have been raised by Mr. Wasmund.

The statement is made in the section entitled "Series-Connected Generators" that propulsion motor field weakening below 40-per-cent field flux is usually not permissible since it can result in bad commutation and also in poor motor performance. Some of the significant limitations that have been pointed out by the author with reference to the series loop system arise directly from this limitation on motor field weakening. I am not an electric machinery designer but, upon some study of the subject, I believe it is not essential to limit motor field weakening to 40 per cent since well-compensated machines can be expected to give entirely satisfactory performance down to much lower values of field strength. With the aforementioned limitation of 40-per-cent field strength removed, the motor associated with a series loop system can accommodate full available engine horsepower under any number of generators in service.

Referring to the section on "Reasons for Selecting D-C Drive," the author is correct in observing that the motor capacity, for vessels such as tugs, ice breakers, and dredges, must be selected with a rating to correspond to the stalled ship propeller characteristics. I believe that this principle applies equally well to ships equipped with series-connected generators and to those equipped with parallel-connected generators. It is not obvious why this is a point to be considered when deciding on the relative merits of series- versus parallel-connected generators as stated by the author.

The information given under the heading of "Stalled Ship Operation" seems to reveal the greatest area of deficiency in operation of the parallel-connected generator scheme. Under the fully stalled ship conditions any reduction in the number of generators in service necessarily involves a reduction in the line voltage and an overload in the generator armature current of the remaining generators if rated horsepower is to be drawn from the generators then remaining in service. If the number of generators is relatively large, this overload current may be acceptable under the desired operating conditions but it nevertheless is a disadvantage of the system. The only way of overcoming this disadvantage is to provide extra flux capacity in the motor over and above that required to meet the stalled ship full horsepower operating conditions and, in order to provide this additional flux capacity, an oversized motor is required.

The author has stated the case well in the concluding portions of his paper wherein he calls attention to conditions that have actually been encountered in a case where five generators were used to supply power to a single armature motor and the low-voltage rating limitation of the series system was not tolerable. In particular I agree with the conclusion of the paper wherein the author observes that for vessels requiring above 2,000 horsepower in single armature propulsion motors and using three or more engines, a study should be made to determine the advantages and disadvantages of the two systems.

Paul G. Tomalin (Alexandria, Va.): Little comment can be made on the author's most interesting paper on the electrical aspects of series and parallel Diesel-electric ship propulsion. He has covered them well and presents a good working basis for analysis.

Some information relative to the mechanical aspects can be offered based on experience with parallel-connected, three-engines-per-shaft installation for ship propulsion. The vessels are ice breakers with the output of three generators applied to a single motor on each of two shafts in the vessel. The engines are rated at 2,000 horsepower and the motor is rated at 5,000 horsepower.

It has been found from experience that the only way in which the main propulsion generators and engines on this class can be properly adjusted to provide full power is to use a water rheostat to load the engines for governor adjustment. The use of the water rheostat allows the adjustments to be made at dock side.

The vessel is pilot-house-controlled by means of pneumatic control to the engine governors. Adjustment of the governors is quite unhandy and the chore of adjustment tedious. Later designs of governors have overcome some of these objections and where parallel systems are to be used particular care should be exercised to assure that the feature of easy adjustment is incorporated in the governors. The governors incorporate a feature known as a torque limiting cam. The purpose of this cam is to limit the amount of fuel that can be injected per stroke as a function of engine speed. The nominal full speed of the engines at full power is 810 rpm and the rating at this speed is 2,000 brake horsepower.

The first step is to adjust the pilot-house control so that no error exists in it so that the same signal and position are obtained on each engine. The engine governor is then adjusted to give 810 rpm at the rated voltage and current. This is done over 3,000 revolutions by counter to get a proper average.

The load is then removed from the engine and a check made on the no-load speed. It has been found that this must lie between 846 and 850 rpm when checked by the same method. This droop is necessary for the proper proportioning of the load between the engines. If the speed droop is not within these limits, adjustments must be made. After these limits are satisfactory the idle speed must be adjusted, in our case to 300 rpm under rated load at that rpm. After this adjustment is all in order, the torque limiting cam is set so that it is just not touching at the full rated load and speed.

These adjustments normally take about three days for six engines and to assure readiness for full power operation at any time they must be accomplished annually.

James A. Wasmund: I am grateful to all of those who presented discussions on the paper. All of the points brought out are well taken and add a great deal to the original conception, which admittedly covered a very narrow field.

Mr. Fifer's point that a fault between the propulsion generator and its setup circuit breaker would probably result in current that would destroy the breaker is well taken. However, since the breaker is located immediately adjacent to the generator terminal box, I believe this risk is very small.

Mr. Coleman's statement that the selection of parallel connection of the generators results in considerable saving in size, cost, and weight of main cable is a pertinent one.

Both Mr. Schaelchlin and Mr. Curry bring out the important point that motorization of a generator will cause reversal of the Diesel engine in a series system and will not in a parallel system. Engine overspeed protection is normally incorporated in the governing system.

Mr. Schaelchlin's comments regarding regeneration during maneuvering are especially appreciated because this question usually comes up at some time during the discussion of every d-c application.

Mr. Curry has pointed out the advantage of parallel connection of generators in providing for possible flooding of motor or engine rooms and still maintaining power on the propeller without making it necessary to enter the flooded compartments. This is especially desirable on an ice breaker.

Mr. Wilde's observation that limitations of the series system are based on an arbitrary minimum field strength of 40 per cent of normal is quite true. We feel that this is justified in a large propulsion motor since the commutating pole flux is increased as the magnetic saturation in the frame becomes less due to less main field strength. This will result in the commutating pole strength becoming too strong at the reduced field speeds for good commutation. It is, of course, possible to arrange to shunt the commutating field winding when the main field is weakened. However, this is usually not desirable on marine propulsion since it does complicate the control.

It is quite correct that a range of constant horsepower by field control for tugs, ice breakers, and dredges is of importance whether series or parallel connection of generators is selected. It does, however, become of more importance in the parallel case when less than the maximum number of generators are in service.

We believe that the tedious and unhandy adjustment of engine governors, mentioned in Mr. Tomalin's discussion, will not be repeated on modern installation. There has been considerable improvement in governor design since the installation mentioned. This, together with generators designed for about 10-per-cent voltage regulation, will provide for proper load division without frequent adjustment.

Fairless Works—Electric Equipment for Slabbing Mill and Blooming Mill

R. H. WRIGHT
FELLOW AIEE

N. L. KINCAID
ASSOCIATE MEMBER AIEE

STEEL ingots produced at the Fairless Works are first reduced for further processing by the most powerful combination of reversing mills which has ever been built. The largest ingots are rolled in a universal slabbing mill and are transferred direct to a hot-strip mill. Smaller ingots are first reduced to rectangular blooms in the slabbing mill and then are transferred to a high-speed 40-inch reversing blooming mill for further reduction to intermediate sizes. This combination has many unique mechanical and electrical features.

Slabbing Mill

Universal slabbing mills have two pairs of plain cylindrical reversing rolls mounted in tandem and spaced so closely that both pairs of rolls work on the metal at the same time. Each set of rolls is driven independently. Two large horizontal rolls are the main rolls and are used for making the heaviest reductions on the metal. Two smaller vertical rolls are used for edging. The separation of each pair of rolls is adjusted by a motor-driven mechanism. Thickness of the bloom or slab is determined by the position of the top horizontal roll and width is controlled by the vertical rolls. A manipulator located in front of the mill is used to turn the ingot on edge so that unusually heavy edging passes can be made in the more powerful horizontal rolls.

MAIN DRIVES

Fig. 1 shows the plan of the universal slabbing mill and roll drives for the Fairless Works. The unusual electrical features of the roll drives are the high capacity of the motors and the scheme of main connections. The horizontal rolls are driven by a 12,000-horsepower (hp) twin-motor arrangement consisting of two 6,000-hp 40/80-rpm direct-connected double-armature reversing motors. One motor drives the bottom roll and the other drives the top roll. These motors are connected to the rolls by universal spindles 30 feet long. The combined maximum torque capacity of the motors is 4,330,000 pounds-feet. The vertical rolls are driven through right-angle reduction gears by a 4,000-hp 60/150-rpm double-armature motor. D-c power for the complete 16,000-hp drive is supplied by a 13,500-kw 8-unit flywheel motor generator which is 83 feet $4^1/_2$ inches long. This set has four 2,500-kw 700-volt generators for the horizontal roll drives, two 1,750-kw 700-volt generators for the edging roll drive, an 8,000-hp 514-rpm wound-rotor induction motor, and a 128,000-pound plate flywheel 14 feet in diameter. Adjustable voltage motors for the mill auxiliaries have a total nominal capacity of 3,900 hp and are supplied by three motor generators having a combined rating of 3,750 kw.

SELECTION OF DRIVES

The largest previous installations of this type have two 5,000-hp motors for the horizontal rolls and a 3,000-hp motor for the edging rolls. While one of these installations has consistently rolled record tonnages, experience has indicated that the mechanical limits have not been reached and that greater output could be obtained if more power were used. For this reason, it was decided to increase the total capacity of the roll drives for the new mill to 16,000 hp. For the last 13,000-hp installation, single-armature construction had been used for all motors. To obtain the 6,000-hp rating required for the horizontal rolls of the new mill without exceeding the required limits of frame diameter, it was necessary to use double-armature construction. The 4,000-hp motor for the edging rolls also was made with two armatures in order to reduce the width of the base and thus facilitate mounting close to the horizontal roll drives.

CONTROL FOR MAIN DRIVES

Control for the roll drives of a universal slabbing mill has four main functions. It permits the operator, by means of a small foot-operated master switch, to control the speed and direction of rotation of both sets of rolls. Also it makes all of the motors operate in unison. Another important function is the automatic control of load division between the motors for the upper and lower horizontal rolls and between the horizontal roll drive as a unit and the drive for the edging rolls. Automatic control of the rate of acceleration and reversal and automatic control of current peaks are also required, regardless of how the master switch is manipulated, so that the operator is free to give all of his attention to the mill.

MAIN CIRCUITS

The arrangement of the main circuits and field circuits for the mill at the Fairless Works is shown schematically in Fig. 2. The main armatures are connected in three separate series loop circuits, each with two motor armatures and two generator armatures. Armatures are sandwiched so that the voltage to ground is that of one armature only. As compared with the usual parallel scheme, this arrangement greatly reduces the possible short-circuit currents, eliminates half of the circuit breakers, and eliminates all but one load balance circuit.

FIELD CIRCUITS

All of the main machines are shunt wound and the field circuits are energized by adjustable voltage exciters. Two exciters are used in a series-sandwich connection in each of the field circuits for the 12,000-hp drive for the horizontal rolls. This arrangement permits the use of small high-speed exciters with quick response characteristics. Each pair of exciters is under the control of a regulating exciter which in turn is controlled, for nonautomatic functions, by the opera-

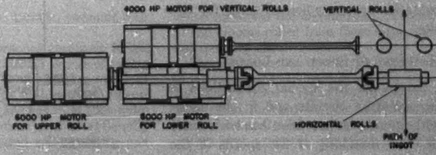

4000 HP MOTOR FOR VERTICAL ROLLS
VERTICAL ROLLS
6000 HP MOTOR FOR UPPER ROLL
6000 HP MOTOR FOR LOWER ROLL
HORIZONTAL ROLLS
PATH OF INGOT

Fig. 1. Plan of slabbing mill and drives

Paper 54-181, recommended by the AIEE Mining and Metal Industry Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 2, 1953.

R. H. WRIGHT is with the Westinghouse Electric Corporation, East Pittsburgh, Pa., and N. L. KINCAID is with the Westinghouse Electric Corporation, Buffalo, N. Y.

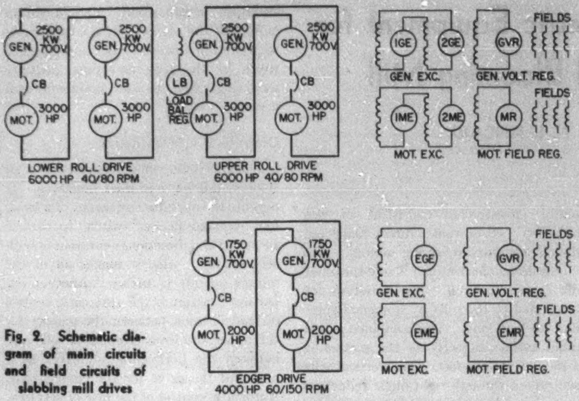


Fig. 2. Schematic diagram of main circuits and field circuits of slabbing mill drives

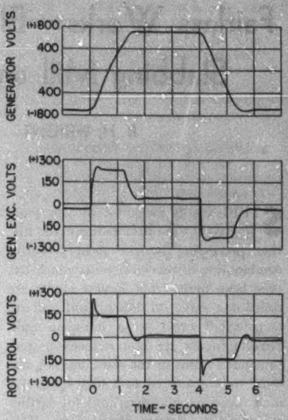


Fig. 3. Response curves of excitation system for generators of slabbing mill when adjusted for 1.5 seconds reversal time

ter's master switch. The regulating exciters have auxiliary fields which provide the required automatic features. A similar excitation system is used for the 4,000-hp edging roll drive except that one exciter is sufficient for each field circuit.

Sequence of Operation

The sequence of the contactors in the master circuit is designed so that all motors have full field excitation when the mill is at rest. When the pedal of the master switch is depressed for either direction of rotation, the voltage of all generators is increased simultaneously. After full voltage has been reached, the field current of all motors is reduced at the maximum rate permitted by the current limit circuits until the desired maximum speed has been reached.

Speed of Response

Excitation systems of this type can be adjusted for a reversal time of 1 second from base speed in one rotation to base speed in the opposite rotation. Practical operating limits are determined by the size and weight of the ingots, the inertia of the mill motors, the inertia of the roll tables on either side of the mill, and the natural reflexes of the operator. Reversal time of 1.5 seconds has been found to be the most satisfactory for the slabbing mill at the Fairless Works. Fig. 3 shows the operation of the generator excitation system when adjusted for this rate.

Matching of Speed and Loads

Except for the special connections in the armature circuits and in the shunt field circuits, no provision has been made for close matching of the speeds of the 6,000-hp motors. This is because rolls of exactly equal diameter may not always be available, thus making it preferable to have equal division of load. The resistance in the armature circuit of each of the motors is sufficient to insure close division of load when the roll diameters are equal. To insure equal division of load under unusual conditions, a load balance system has been provided which balances loads by differential adjustment of the voltage in the loop circuits.

The heavy sections which are rolled in a large slabbing mill can transfer a considerable amount of power between the main rolls and the edging rolls. With the large difference in the capacity of the main roll drive and the edging roll drive, the edger and its drive are more likely to be damaged than is the metal if the work is not proportioned properly. To make sure that the edger carries no more than its share of the total load, the edger drive has a drooping speed characteristic. Two rheostats are also provided in the field control for the edger motor for adjusting the speed relations. One rheostat is calibrated for the diameter of the main rolls and the other for the diameter of the edging rolls. In addition to these features, the control provides for a slightly lower speed for the edger motor when the ingot is moving from the edging rolls to the main rolls to compensate for the reduction in the main rolls.

Slabbing Mill Auxiliaries

Ingot Cars

Two ingot cars, operating on closely spaced parallel tracks, are used to carry

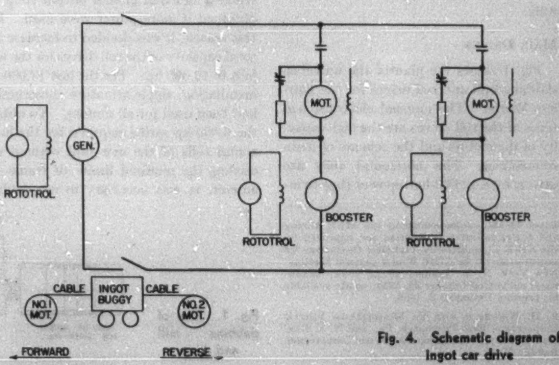


Fig. 4. Schematic diagram of ingot car drive

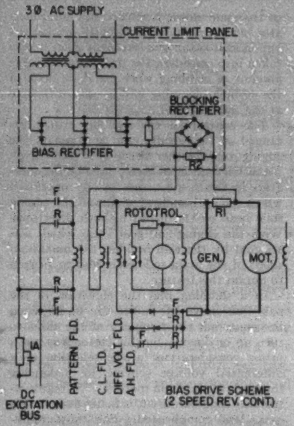


Fig. 5. Schematic diagram of control for adjustable voltage auxiliaries

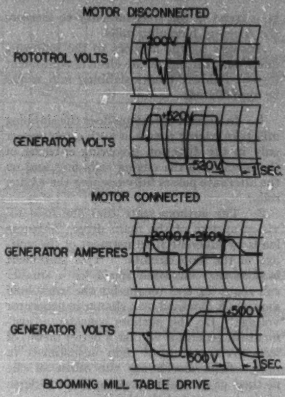


Fig. 6. Performance curves of a typical adjustable voltage drive

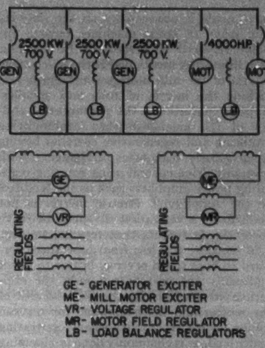


Fig. 7. Main connections for 8,000-hp drive for blooming mill

the ingots from the soaking pits to the entry side of the mill. The close spacing of the tracks made it practically impossible to install collector systems for propulsion motors on the cars so they are driven by cables. The unique feature, used here for the first time, is that each car has a cable at each end with a motor-operated drum at each end of the track. Both cables are held under tension at all times for either direction of travel to prevent contact of the cables with the track. A schematic diagram of the electric connections is shown in Fig. 4. This propulsion system is designed for the unusually high speed of 1,160 feet per minute. Automatic positioning systems with push-button controls are used to spot the cars for loading at the soaking pits.

Mill Auxiliaries

Fig. 5 shows the basic circuits for the adjustable voltage drives for the mill tables, manipulators, horizontal roll screwdown, and vertical roll adjustment. The main control element is a high-gain voltage regulator having a current limit feedback circuit. This regulating system causes the generator to operate at approximately 200-per-cent voltage at light load and limits the maximum current at 250 per cent. These characteristics provide the maximum rate of acceleration and deceleration without damage to the motor. Performance of a typical application is shown in Fig. 6.

An unusual arrangement has been used for adjusting the position of the vertical rolls. The usual practice is to connect the two mechanisms mechanically or to synchronize them electrically to insure that both rolls are always the same distance from the center of the mill. In this installation there is no mechanical or electrical tie between the two mechanisms, each being driven by an independent, adjustable voltage motor. The motors operate in unison when the rolls are brought together or separated. A differential electric indicator shows when either roll is out of alignment. It has been found that the master control for this system can be adjusted so closely that the errors in alignment are negligible.

Blooming Mill

Reversing blooming mills have two horizontal rolls which are cast with matched collars which form shallow rectangular openings for shaping the metal. The upper roll can be raised and lowered, and the separation of the rolls and the roll opening used determine the dimensions of the product. A manipulator in front of the mill is used to turn the ingot and to guide it into the desired roll opening.

The 40-inch blooming mill for the Fairless Works has a conventional 8,000-hp twin-motor drive consisting of two 4,000-hp 65/130-rpm reversing motors supplied by a 7,500-kw flywheel motor generator. This is the greatest continuous capacity which has been used for a secondary blooming mill and the speed of 65/130 rpm is unusually high for a reversing drive of such capacity. This combination of power and speed was selected to insure fast rolling of blooms as they are received from the universal mill. The drive also has capacity to roll direct from ingots the largest sections which the mill can handle mechanically. Fig. 7 shows the main connections.

This combination of mills is designed to roll a wide variety of slabs and blooms. Slabs 6 inches to 8 inches thick up to 74 inches wide and weighing up to 40,000 pounds are delivered direct from the slabbing mill. Blooms of all sizes are rolled by passing the steel through both mills. The installation has exceptional tonnage possibilities.

◆

Discussion

H. E. Larson (General Electric Company, Schenectady, N. Y.): This short paper which describes the mechanical and electrical features of the world's largest slabbing mill equipment, as well as the 40-inch companion blooming mill, is indeed, to the authors' credit, concise, laconic, succinct, and to the point.

Deserving of additional emphasis is the unusual practice at Fairless Works of rolling blooms from ingots in a slabbing mill and a companion blooming mill, without reheating, which no doubt springs from economies gained in casting fewer sizes of exceptionally large ingots. Increased rolling of the hot metal improves the metallurgical properties of the blooms rolled from these large ingots and reduces sheared end losses.

The authors state that excitation circuits of a reversing mill main drive can be adjusted for a drive reversal time of 1 second—base speed forward to base speed reverse. It is of interest to note the authors' observation that practical operating limitations have made it unnecessary on the slabbing mill drive to stress machines beyond the point required to reverse in 1.5 seconds. One limitation mentioned is the mechanical inertia of the mill motors which must also include the inertia of rolls and spindles. Another limitation is the commutating ability of the machines when commutating motoring and regenerative currents having tremendously high rates of change. With reversing drive control using rotating regulating power amplifiers such as the Rototrol or the Amplidyne, electric inertia is not necessarily a limitation of speed of drive reversal. It follows that reversal in, say, 1.20 seconds, and less from base speed forward to base speed reverse, begins to affect materially the cost of not only the excitation circuits but also the size and cost of the motors and the generators for a given rolling torque availability when accelerating under load, as is commonly done in blooming mill practice.

The sandwich, or alternate series connection of motor and generator armatures in a single loop, is a well-known simple connection offering many advantages though not too often used in the United States. One reason is that the parallel connection described by the authors for the main drive of the 40-inch blooming mill permits full speed operation at reduced torque in the event of trouble with one generator. However, since it is quite likely that the seven 2,500-kw 700-volt generators of the Fairless installation are duplicates and protected by a spare armature at least, this objection to the use of the series circuit on the slabbing mill reduces somewhat in importance.

The authors correctly cite the 65/130-rpm speed rating of the 40-inch blooming mill drive as being unusually high. A new 44-inch blooming mill twin drive rated 12,000 hp recently placed in operation in the Pittsburgh area and establishing new records uses two double-armature 6,000-hp 70/140-rpm motors which are even higher in speed rating than those of the 40-inch blooming mill drive at Fairless Works. If two examples are sufficient evidence, then there seems to be a trend in the steel industry to increase the speed of blooming mills.

Answers by the authors to the following questions relating to the electric equipment for the Fairless Works slabbing mill would be appreciated:

1. In normal operation does the slabbing mill take edging passes in both directions or only in the direction involving entering of the metal into the edging rolls first and on the alternate passes fully opening the edging rolls?
2. The authors state that the load division system of the main drive "balances loads by differential adjustment of the voltage in the loop circuits." This statement implies that increased speed for a smaller roll requiring less torque for the same load amperes is obtained by a change in generator voltage. Do the authors consider adjustments in generator voltage a better means of load division control than adjustment in motor field considering the effect of the method of control in turn-up or turn-down of the front end of the slab?
3. Conventional reversing mill practice would call for a 10,000-hp wound-rotor induction motor to drive the flywheel motor-generator set. Are the authors convinced that this practice is too conservative since they have selected an 8,000-hp motor?
4. Are the "drooping characteristics" of the edger drive obtained by current limit control using edger motor field strengthening or edger generator voltage reduction?
5. In determining 1,160 feet per minute as the maximum speed of the ingot car, what was the maximum length of track and the maximum number of round trips per hour used in the calculations?

R. H. Wright and **N. L. Kincaid:** After considerable deliberation the series-sandwich system of main connections was used for the slabbing mill drives because of its simplicity. It was felt that the simplicity would outweigh any advantages of a parallel arrangement. Armature connections are arranged so that one motor armature and one generator armature in each circuit can be cut out in the event of a failure. Also it was taken into consideration that a considerable amount of slabbing work could be done in the 40-inch blooming mill in case of trouble at the slabbing mill.

The control for the slabbing mill is arranged so that drafts can be taken in the edging rolls in either direction of slab travel. No restrictions have been placed on the operators in this respect and for any operation which the authors have observed the edging rolls bear on the metal in both directions. With the edging roll gear ratio which was selected by the mill builder, the maximum speed of the main rolls is limited to 70 rpm to obtain this feature.

In a slabbing mill the elevation of the pass line above the tables can be adjusted to counteract the tendency of the slabs to turn up or down. So this factor was not a prime consideration in the selection of a differential voltage scheme for balancing the loads of the main motors. Voltage control was used because it is faster and, in this case, less complicated. For most applications it is necessary and desirable to operate on the motor fields.

In selecting the driving motors for the flywheel sets, the slabbing mill was considered as a special case. On the basis of the small elongation in the mill and low energy requirement per ton of product, there was no justification for using a motor larger than 8,000 hp. Under other conditions it is often desirable to use a motor which is larger than the usual standard rating.

Drooping characteristics of the edger drive are obtained by a combination of motor field control and voltage control. For most conditions the edger motor has sufficient field range for field control.

The track for the ingot cars is about 650 feet long. There are ten soaking pits and the no. 1 pit is nearest the mill. The maximum speed of the cars was based on making a round trip of 900 feet from pit 7 in the time required to roll one ingot, which is about 1 minute.

A World-Wide Standard Traction Motor for Diesel-Electric Locomotives

M. J. BALDWIN
MEMBER AIEE

Synopsis: Dieselization of steam railroads abroad has recently been proceeding at a greatly accelerated pace. This has served to fix attention on the need for a motor suitable for general freight and passenger service on track gauges down to 1 meter, and with axle loads limited to 44,000 pounds per axle. A motor designed for such applications was placed in production last year. Its characteristics give a locomotive performance, within these limitations, comparable to that of many domestic units. Surveys indicate that this covers the needs of a great majority of narrow-gauge operations.

World Survey

TRACTION motors for Diesel-electric transportation applications are so standardized today that the appearance of a new one is a noteworthy event in the history of the industry. Such a motor, designated as the *761*, was put into mass production by one manufacturer during 1953. Its inception was due to the fact that dieselization of steam roads is beginning to make rapid progress throughout the world. Numbers of lines in Australia, Asia, Africa, and South America are now converting from steam to Diesel power. Still others are considering making the change.

In North America practically all the large roads use a traction motor of similar design, designated as the *752*, on at least a portion of their Diesel-electric locomotives. Its performance record is

---

Paper 54-189, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE North Eastern District Meeting, Schenectady, N. Y., May 5–7, 1954. Manuscript submitted October 21, 1953; made available for printing March 1, 1954.

M. J. BALDWIN is with the General Electric Company, Erie, Pa.

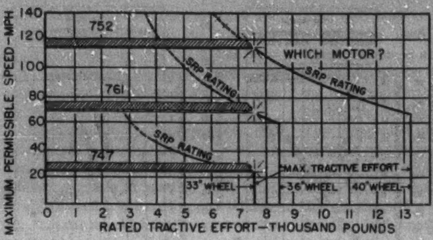


Fig. 1. Comparison of application capacities of standard design traction motors

well known, and it has become one of the standard motors of the railway industry. Outside of North America, however, a smaller motor appears to be necessary except for possible applications on a few of the larger railroads in Latin America.

There are several reasons for the need of a smaller motor in foreign service. Economic or strategic considerations, for example, frequently dictate the use of lightweight rails. Rails weighing 60 to 80 pounds per yard are quite common throughout the world. For the same reasons bridge structures are generally lighter than those found on American railroads. The motor here described was designed to meet the needs of a large percentage of these foreign roads. Hence it is being offered for world-wide application in much the same way as the *752* now serves in the domestic market.

Some railroads, such as the Paulista in Brazil and the state railways of Chile, could possibly use the larger domestic motor. Even on these roads, however, the new motor is definitely a competitor, since they cannot effectively use anywhere near the 70,000-pound maximum axle loading on which the domestic motor may be applied.

Market surveys indicate that at least 70 per cent of the present demand in foreign business is for a motor good for loadings of from 25,000 to 35,000 pounds per axle. The new motor was designed light enough to be applicable in this weight range when needed. If it is to be a standard motor, however, it must have the broadest possible field of application. Therefore, with future growth in mind, it has been designed with capacity for any axle weight up to 44,000 pounds. Estimates indicate that only 10 per cent of all foreign railroads use axle loads heavier than this.

One of the fundamental, and probably most important, reasons for designing a new motor is the need for a motor good for road-service speeds on narrow-gauge track. Domestic motors, with the exception of the relatively small-capacity double-reduction geared type, designated as the *747*, are designed for standard-gauge track. Both 42-inch gauges and meter gauges are extensively used in Asia, Africa, South America, and Australia. For this reason, the new motor has been designed so that it can be used on any gauge down to 1 meter in width. This should make it applicable to 95 per cent of all potential applications within its axle weight range.

From the foregoing, it is evident that the two basic reasons for designing a new motor were to obtain one which more nearly meets the weight-per-axle limitation prevalent on foreign railroads, and to obtain a motor which is applicable on narrow- as well as wide-gauge lines.

As has already been mentioned, the double-reduction geared motor is applicable on gauges down to a minimum of 36 inches. It is a standard motor widely used both in this country and abroad on applications where high speed is not essential. The choice of double reduction as compared to single is determined principally by the commutator surface speed. For service conditions today this is generally less than 11,000 feet per minute. It is economical to use double reduction where maximum service speed of the locomotive is low enough, and the commutator diameter small enough, to meet this limitation in peripheral speed. Such is the case for the *747*, but not for the *761*. The new motor, being single-reduction, for this reason more nearly resembles the *752*, and hence is compared with that motor in this paper.

Application Capacity

Fig. 1 has been designed to compare the application capacities of the three motors just discussed. The horizontal scale represents continuous tractive effort, or train-hauling capacity, of a single motor. One such particular value of train-hauling capacity is indicated by the small train for each of the three motors. The elevation, or track level, of each train is equal to its maximum permissible speed. All three trains are assumed to require equal continuous tractive effort. The domestic motor, which weighs 7,850 pounds, will pull the train with a maximum permissible speed of 115 miles per hour (mph). A motor of this size is required only if high top speed is needed. The double-reduction motor, which weighs 2,985 pounds, will do the same job with a maximum permissible speed of 25 mph. This is suitable if the service is such that low speed only is required. The new motor finds its place midway between these two. It weighs 4,250 pounds, and will haul the same sized train with a maximum permissible speed of 70 mph.

Suppose the maximum speeds, or elevations, of the three trains in Fig. 1 had been made equal. For example, assume a railroad where, because of terrain, track and gauge limitation, the maximum permissible speed is 40 mph. The continuous tractive efforts available from the three motors would be as shown in Table I. It is evident that, whether the comparison is made on the basis of identical tractive effort or identical speed, the new motor admirably fills the gap between the two older standard motors.

In compiling Fig. 1 no attempt was made to indicate the various standard combinations of wheel size and gear ratios obtainable. The minimum wheel diameter for each motor is indicated. Also, for that minimum wheel size, the continuous tractive effort corresponding to maximum gear reduction is shown as the right-hand boundary line for each motor area. The field of possible application

Table I. Comparison of Continuous Tractive Efforts at 40 MPH

Motor	Continuous Tractive Effort, Pounds
Double reduction	4,650
New design	8,500
Domestic	13,250

Table II. Comparison of Motors on Basis of Maximum Gear Reduction

Motor	Maximum MPH	Continuous Tractive Effort, Pounds	Continuous Adhesion, Per Cent	Wheel Diameter	Maximum Axle Weight, Pounds
Domestic	65	13,250	18.9	40	70,000
New design	60	8,500	19.3	36	44,000
Double reduction	25	7,600	15.2	33	50,000

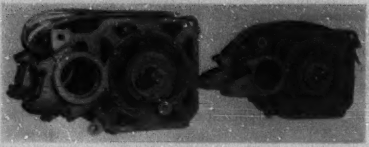


Fig. 2 (above). Comparative size of the domestic (left) and world-wide (right) traction motors

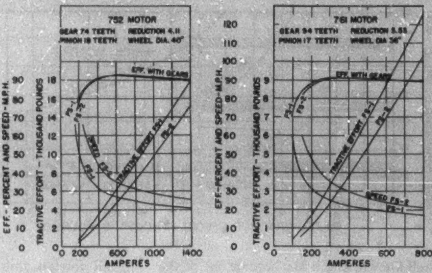


Fig. 3 (right). Characteristic curves for the domestic (752) and world-wide (761) motors operating on 600 volts

of the three motors can be derived from the information set forth in Fig. 1. In view of the fact that most applications are made on the basis of maximum gear reduction, the information given in Table II may prove useful.

It is perhaps too early to predict the size of the locomotive unit for which the new motor will prove most useful. In the United States, the 4-motor, 1,600-horsepower (hp) Diesel-electric locomotive unit is most generally used today. An examination of Table II shows that a unit with six of the new motors, giving a total continuous tractive effort of 51,000 pounds, is closely comparable to a unit with four domestic-type motors, giving a total continuous effort of 53,000 pounds. This means that a locomotive for 60 mph maximum speed and equal service may be practically as heavy, and hence haul almost as heavy a train, when equipped with six new motors as with four domestic motors. In addition to this, the characteristics of the two motors make possible the use of the same basic power plant, so that it would seem that one of the locomotive units for foreign service might well be a 1,600-hp 6-motor one. Most of the applications made so far have been on this basis.

Other applications of a 6- motor unit at 1,200 and 1,800-hp are also probable. Branch lines in North America laid with lightweight rail may find such a locomotive useful when ballasted up to the maximum weight of 132 tons on drivers. On application where the 6-motor locomotive might prove uneconomically large, the new motor may be used on 4-motor locomotives of 660, 800, 1,000, or 1,200 hp with any weight up to 88 tons on drivers.

The locomotive weights mentioned are considered a ballasted maximum for the maximum reduction 60-mph gearing. When higher maximum speeds are required, the permissible weight will be somewhat less. Most applications, however, are expected to be limited by rail weight and permissible bridge loading to 25,000 to 35,000 pounds per axle. The new motor should, therefore, have ample margin to serve in the capacity of a standard motor and to cover adequately any application in its field.

Table III. Condensed Motor Specifications

	Domestic	World-Wide
Number of exciting poles	4	4
Number of interpoles	4	4
Present insulations	mica glass, silicones, and permafils	
Number of brush holders	4	4
Brushes per holder	3	2
Frame type	box	box
Bearings		
Armature	antifriction, grease lubricated	
Axle lining	sleeve, oil lubricated with felt wick	
Spider	none	none
Armature winding	multiple	multiple
Field winding	series	series
Ventilation	blown	blown
Gear covers	fabricated steel, lap sealed at joints	
Gear lubrication	gear compound	gear compound

Comparison of Design

A comparison of the relative size of the present domestic motor and the new world-wide motor is given by Fig. 2. It also shows clearly the general similarity in design of the two motors. This similarity is further indicated by the condensed specifications shown in Table III. The characteristic curves of the two motors are shown side by side in Fig. 3 for purposes of comparison.

While the maximum permissible speed of the new motor is 3,100 rpm as against 2,280 rpm for the domestic motor, the diameter of the commutator has been reduced sufficiently to keep the peripheral velocity within 6 per cent as low as that of the domestic motor.

Commutation is not the problem in the new motor that it is in the larger *752*. It is a known fact in motor design that, all else being equal, the larger the armature diameter, the greater the hp a motor will handle from the standpoint of commutation. The *761* has 93 per cent of the armature diameter of the *752*, yet it is required to transmit a maximum of only 300 engine hp, or about 54 per cent of the 560 engine hp that the latter normally transmits on Diesel-electric passenger locomotives. Despite this difference the new motor retains the multiple type of armature winding.

Like its predecessor, the new motor has an arch-bound rather than a V-bound commutator. In the arch-bound construction used on these more modern motors, the V-shaped sections of the cap and shell merely draw the bars in toward the center. This builds up tangential or arch pressure between the segments. In

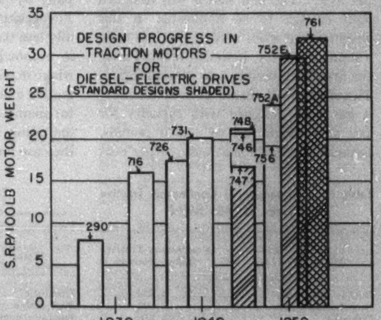


Fig. 4. Graphic representation of progress in design of traction motors for Diesel-electric locomotives

the V-bound construction the bars are clamped axially between the V-shaped rings with relatively low bar-to-bar pressure. Both types, if properly cured, are good but the arch-bound generally has the better surface.

The commutator of the new motor, like that of the domestic motor, is cured by the "spin" process. This consists of a number of cycles of heating while being driven at high speed, cooling and tightening, that are repeated until a commutator surface is obtained that is smooth and relatively unaffected by variations in either speed or temperature. Considering all factors, the inherent electric design constants, the type of winding, the brushes, and the commutator construction, the new motor should have a record of commutator maintenance and brush life at least comparable to that of the domestic machine. The principal reason for this is the relatively low power transmitted per inch of armature diameter.

Although insulating materials are continually being improved, the insulations used in these motors are quite similar. Wherever possible, organic materials have been taken out and inorganic materials substituted. The basic materials are mica and glass. Bonding material is principally silicone resin with protective permafil coatings where needed. Comminuted mica in the form of mica mat is used in the armature of the new motor.

Both motors have the armature windings "reroll banded" into the slots. In this process a layer of temporary binding wire, with both ends anchored after application, is used. One free loop, to which tension is applied by means of a pulley, distributes the pressure by traveling back and forth from one end of the binding to the other as the core is rotated. This is repeated until no more slack can be taken up by the traveling loop. This process, done while the winding is hot, forces the coils down solidly into the slot and on the end windings in preparation for the final banding of the ends and wedging of the core.

Design Progress

The relative elevations of the trains in Fig. 1 indicates the relationship between the motor application capacities. The elevation is proportional to the speed ratio power (SRP)[1] of each motor

$$\text{SRP}=\frac{\text{maximum speed}}{\text{rated speed}}\times\text{rated hp}$$

This can be written in the following form

$$\text{SRP}=\frac{\text{maximum speed in mph}\times\text{rated tractive effort}}{375}$$

In these equations rated tractive effort is the maximum which can be sustained continuously by the motor. The corresponding speed is the rated speed, and the corresponding motor output hp is the rated hp.

In the latter form the expression indicates clearly that if the continuous tractive efforts are made equal by the use of suitable gearing, the maximum speed is proportional to SRP. Conversely, if the maximum speeds are made equal, the continuous tractive effort is proportional to SRP.

To show approximately the design progress, Fig. 4 gives a comparison of SRP values per hundred pounds of motor weight for each of the three motors plotted against year of design. Also shown are a number of other motors, some of which have been manufactured in considerable quantities but which have never been adopted as standard design. This comparison indicates the advanced design of the world-wide motor.

Conclusion

Dieselization of steam railroads abroad stands today in approximately the same position it occupied 10 years ago in the United States. This means that an era of great activity lies immediately ahead. The fact that there is in existence a universal standard motor for road passenger and freight applications has great technical and economic significance. This new motor should help to promote the rapid spread of Dieselization abroad.

Reference

1. Twenty-Five Years Progress in the Design of Traction Motors, M. J. Baldwin. *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 182–87.

◆

No Discussion

Transient Analysis of Sampled-Data Control Systems

G. W. JOHNSON
ASSOCIATE MEMBER AIEE

D. P. LINDORFF
ASSOCIATE MEMBER AIEE

RECENTLY, studies of sampled-data control systems have been published[1,2] which furnish exact and approximate expressions for the closed-loop transient response to an arbitrary external disturbance. However, the computations required are exceedingly tedious, even for the simplest type of system. Alternatively, a locus plotted in the frequency domain has been proposed by Linvill,[1] to permit studies to be made of closed-loop stability (in accordance with Nyquist's criterion). Here again the computations are tedious. It is further pointed out that the Nyquist plot obtained in this manner applies only at the sampling instants and, as such, the well-known techniques pertaining to relative stability of continuous systems do not readily apply, although absolute stability is well defined by this technique. Consequently, in most practical situations, the designer has found it expedient to resort to analogue computer studies to determine how the transient response depends upon such factors as the sampling frequency or the velocity constant of the proposed system. The purpose of this paper is to furnish a set of dimensionless curves, based upon computer studies, which are of a sufficiently generalized nature to permit approximate design criteria to be established for a wide variety of system transfer functions.

Fig. 1 shows, in basic form, the feedback-control system of interest. $G(s)$ represents the parts of the system described by conventional linear transfer functions, together with such time delay as would

---

Paper 54-196, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE North Eastern District Meeting, Schenectady, N. Y., May 5–7, 1954. Manuscript submitted January 27, 1954; made available for printing March 18, 1954.

G. W. Johnson and D. P. Lindorff are with the University of Connecticut, Storrs, Conn.

The research described in this paper was done at the University of Connecticut under the sponsorship of the International Business Machines Corporation in co-operation with the United States Air Force. The authors are especially grateful for the helpful suggestions made by Dr. R. J. Kochenburger and Dr. C. G. A. Nordling, and also wish to acknowledge the aid and co-operation of other co-workers on the International Business Machines project.

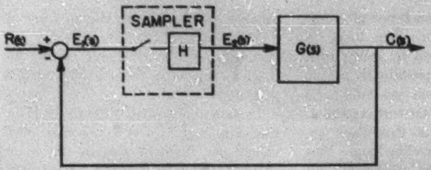


Fig. 1. Basic sampled-data system

arise from a digitally computed source of error. The sampler simulates the portion of the system responsible for the sampling process, including a hold circuit which is normally employed in sampled-data systems. The sampler can be represented physically as consisting essentially of two parts:

1. A switch which closes momentarily at each sampling instant (the time during which the switch is closed must be short compared to a sampling period).
2. A holding circuit H, which may be looked upon as a low-pass filter.

In this paper, a zero-order hold system[2] (or "clamper") is used. The clamper circuit was chosen since it represents the most commonly used type of holding device. Its use is justified in a majority of instances because of its simplicity and its desirable filter characteristics.

Description of Method

A brief statement of the techniques to be developed later will be presented to clarify the method used to predict the approximate transient response of a sampled-data control system.

It will be shown that a system having an open-loop transfer function defined by

$$G(s)=\frac{K_v}{s(\tau s+1)^n} \tag{1}$$

can be used to approximate the transient response of any control system of the linear, type-1 variety that has not been partially equalized by lag-network compensation.[4] Since the system defined by the transfer function in equation 1 is to approximate the actual system of interest, it will be referred to as a pseudo system. In any particular problem, the pseudo-system transfer function is specified by selecting the appropriate values for τ and n in equation 1 so that the two systems have nearly equal open-loop transient responses. If two systems have nearly equal open-loop responses, their closed-loop responses will be similar regardless of whether the error is supplied continuously or in sampled-data form.

The pseudo system can be written in dimensionless form as follows

$$G(w)=\frac{K_w}{w(w+1)^n} \tag{2}$$

where

$w=\tau s$
$K_w=K_v\tau$

or in terms of frequency response

$$G(ju)=\frac{K_u}{ju(ju+1)^n} \tag{3}$$

where

$u=\tau\omega$
$K_u=K_v\tau=K_w$

Analogue computer studies were made on the dimensionless pseudo system for values of $n=1$, 2, 3 and 4, and the results were plotted as a function of sampling period in Figs. 2, 3, and 4. A 20-per-cent peak overshoot in step-function response was used as a criterion of stability.

The following procedure should be followed to analyze the effects of various sampling rates on a control system under consideration:

1. Select a pseudo system that has approximately the same open-loop transient response as the control system under consideration. Time delay may exist in the open loop of the actual system.
2. From the curves in Fig. 2, predict the velocity constant for a 20-per-cent peak overshoot in step function response for the value of sampling rate of interest.
3. From the curves in Figs. 3 and 4, determine the rise time and response time for the

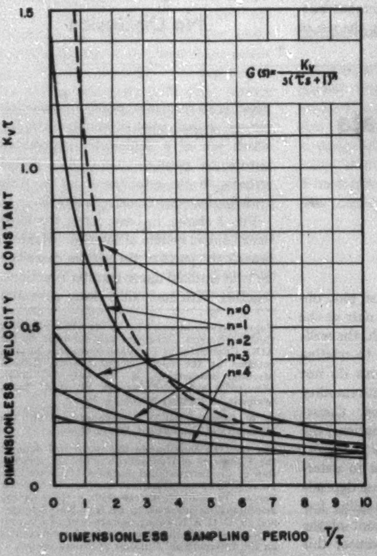


Fig. 2 (left). Dimensionless plots of velocity constant versus sampling period for the pseudo system based on a 20-per-cent overshoot in step-function response

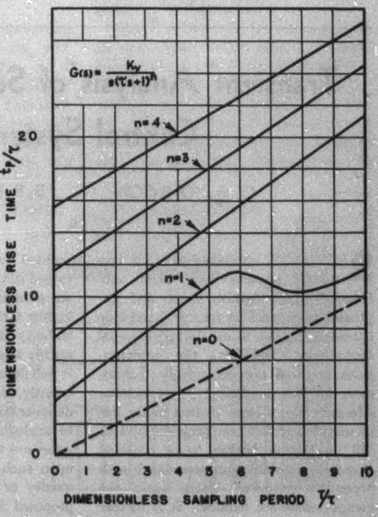


Fig. 3 (right). Dimensionless plots of rise time versus sampling period for the pseudo system based on 20-per-cent overshoot in step-function response. t_p is measured from the instant when the error is first sampled

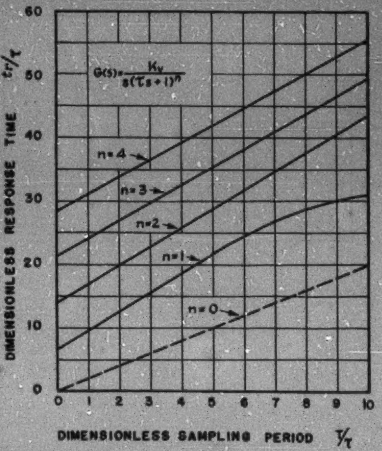


Fig. 4. Dimensionless plots of response time versus sampling period for the pseudo system based on 20-per-cent overshoot in step-function response. t_r is measured from the instant when the error is first sampled

sampling period just mentioned, to obtain a complete picture of the step function response.

This technique will be explained in detail and illustrated with specific examples.

Selection of the Pseudo System

A pseudo system of the form

$$G(s)=\frac{K_v}{s(\tau s+1)^n} \tag{1}$$

can be made to have a continuous transient response similar to an important category of control systems by making the proper selection of τ and n. The method used to accomplish this will be developed on the basis of matching the real parts of the open-loop transfer functions. Removal of the integration is expedient since, subject to certain restrictions, the transient response of a system that contains no poles in the right-half s plane, or on the imaginary axis, is uniquely related to the real part of the system transfer function with $s=j\omega$. This relationship is developed in Appendix I, where its application to the pseudo system, including the case of noninteger values of n and time delay, is justified.

Fig. 1 shows the basic block diagram of a totally equalized type-1 linear-control system wherein the error is being supplied in sampled-data form and clamped between sampling instants at the previously sampled value. Since the continuous portion of the open-loop transfer function $G(s)$ contains a pole at $s=0$, the real part of $j\omega G(j\omega)$ must be considered if the restrictions are to be met in relating the real part to the transient response. Upon removal of the integration, $sG(s)$ does not contain poles in the right-half s plane or on the imaginary axis if the range of application is restricted to systems that are open-loop stable.

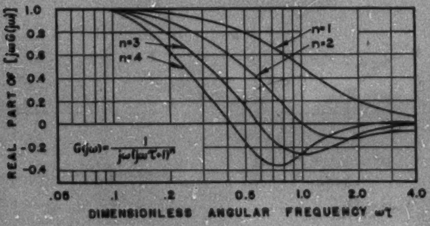
It will be shown later, in example 1, that the most desirable pseudo system may require the use of a noninteger n. Although a pseudo system with noninteger n is not a ratio of rational polynomials in s, it is shown in Appendix I that its transient response is still analytically related to the real part as a function of ω. Therefore, mathematically, its transient response does exist and a pseudo system of this form can be used to predict the transient response of an actual system that has a similar real part.

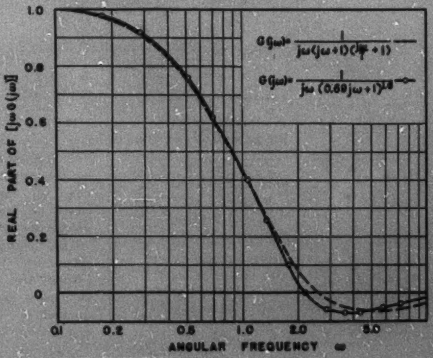
Similarly, the transient response can be related to the real part for systems containing time delay within the major loop, wherein the time delay can be incorporated within the open-loop transfer function in the form of $G'(s)=G(s)\epsilon^{-as}$. If the pole at $s=0$ is removed, we can plot the $Re\ j\omega G'(j\omega)$ and select a pseudo system, without time delay, by matching real parts so that the open-loop pseudo system response is similar to the actual system response including time delay.

Since the real-part plot of a given transfer function completely specifies the transient response (subject to limitations as discussed in Appendix I), the appropriate pseudo system can be considered as one that best matches the real-part plot of the actual system. It might at first appear immaterial as to whether the pseudo system is selected on the basis of matching the real parts defined by the open-loop or the closed-loop systems. However, it is to be pointed out that there are advantages to be gained by maching on an open-loop basis. The most significant advantage in working directly with the open-loop transfer functions becomes apparent when one considers the effects of sampling the error. For sampling, in

Fig. 5 (below). Dimensionless plots of modified open-loop real parts for the pseudo system

Fig. 6 (right). Real-part plots to show the correlation between the actual and pseudo systems for example 1





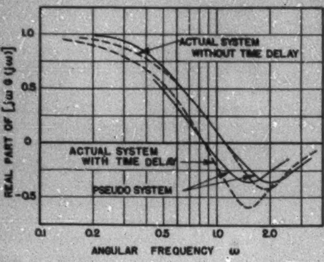
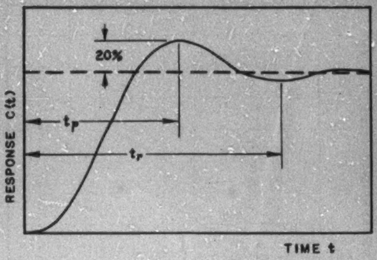


Fig. 7 (left). Real-part plots to show the correlation between the actual and pseudo systems for example 2

Fig. 9 (right). Typical step-function response curve of an under-damped control system



itself, places emphasis on the open-loop response of the system. Furthermore, the transient response of the open-loop transfer function $G(s)$ is related to the transient response of the modified transfer function $sG(s)$ by a time integration. Such an integration process tends to smooth out the discrepancies in the open-loop transient response derived from the real-part plots of the modified transfer function. Hence, larger discrepancies between the actual and pseudo system real-part plots can be tolerated than would be the case if an integration process were not involved.

Since the real-part plot is uniquely related to the transfer function from which it is derived, the criterion for the selection of the pseudo system will be based upon what compromise in matching real parts leads to the most favorable agreement in the open-loop transient responses. To accomplish this, it is helpful to recognize certain features that characterize the real-part plots of transfer functions under consideration, as follows:

1. All curves have a magnitude of unity at zero frequency (for unity K_v).

2. The peak value of the negative real part depends upon the distribution of the poles of the open-loop transfer function.

3. For the pseudo system, the frequency at which the real part first changes sign can be shifted by changing the value of τ without affecting the magnitude of the negative peak.

In the case of the pseudo system, it can be seen, with reference to Fig. 5, that by permitting n to assume noninteger values, the magnitude of the negative peak will take on a continuous range of values from zero (for $n=1$) to 0.35 (for $n=4$). Thus by selection of the appropriate value for n, together with the proper choice of τ to bring about the desired shift of the real part on the dimensional frequency axis, the satisfactory pseudo system can be selected. Certain limitations are imposed as will be discussed later.

The point of crossover of the real part (where the real-part plot first assumes a zero value) is important in effecting the proper time base for the open-loop step-function response.[3] Hence, a first approximation for the selection of τ in the pseudo system is established. Therefore, the following general criteria for matching real parts are established:

1. Select the value of n in equation 1 which best matches the magnitudes of the negative peaks.

2. Select a value of τ in equation 1 to match the crossover frequencies. (A slight modification of this value of τ may be desirable, as pointed out later in example 1.)

Consideration of the basic form of the pseudo system serves to point out certain of its limitations. As n approaches ∞, it can be shown that the magnitude of the negative peak in the real part approaches unity. However, for transfer functions containing only real roots, the negative peak of the real part is usually small enough in magnitude to be matched by a pseudo system having $n \leqslant 4$. See Fig. 6 for a specific example. It is only

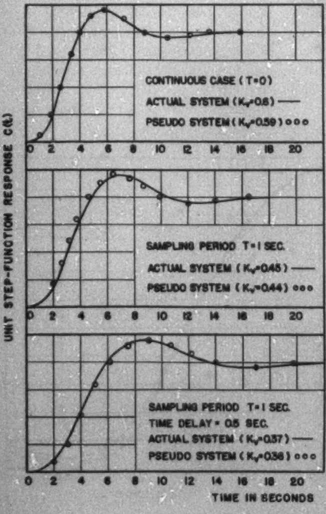
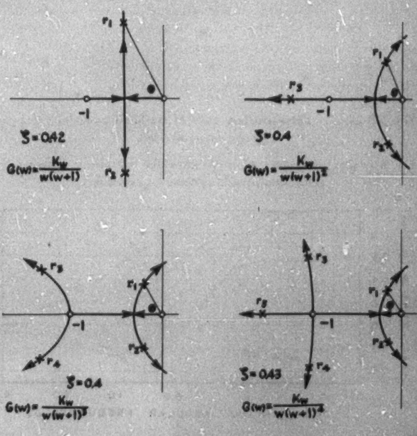


Fig. 8 (left). Transient response curves for the actual and pseudo systems to show the correlation of results for example 2

Fig. 10 (right). Root loci showing the distribution of the closed-loop poles for the continuous pseudo system with K_v adjusted for a 20-per-cent overshoot in step-function response



when the open-loop transfer function contains a pair of significant complex roots associated with a small damping ratio ζ that the negative peak of the real part may become abnormally large. This case will be illustrated in example 2, and is shown graphically in Fig. 7. One might expect that a pseudo system with $n>4$ would handle this case. However, it has been found that for highly underdamped systems, the over-all contour of the real part of the pseudo system falls short of matching the actual system even though the magnitudes of the negative peaks are identical. In other words, for underdamped systems, a pseudo system with $n=4$ represents as good a match of real parts as one can obtain, though it will be shown in example 2 that sufficiently accurate results can be expected for this case.

The following examples will serve to illustrate the pseudo system approach to the analysis of sampled-data systems including the effects of time delay within the major loop. The continuous portion of the control systems are of the linear type-1 variety, and have not been partially equalized with lag-network compensation.

EXAMPLE 1—OPEN-LOOP TRANSFER FUNCTION CONTAINING REAL ROOTS

To illustrate the use of a pseudo system characterized by a noninteger n, consider a system transfer function defined by

$$G(s)=\frac{K_v}{s(s+1)(s/7+1)} \tag{4}$$

A plot of the real part of $j\omega G(j\omega)$ for $K_v=1$ is shown in Fig. 6. To match the magnitude of the negative peak for this case, a pseudo system having $n=1.6$ was chosen. It was seen that, for the pseudo system to have the same frequency of crossover as does the original system, τ must be equal to 0.57. However, the need for a slight adjustment in this value of τ became evident when the real-part plots of the actual and pseudo systems were superimposed, because it is more important that the low-frequency portions of the two curves match if time of rise is to be faithfully approximated. When plotted on a logarithmic frequency scale, a change in τ results in a simple shift of the real-part plot. To minimize errors between the curves, a τ of 0.69 was selected, and the real part of the pseudo system was plotted for this value of τ in Fig. 6.

With reference to Fig. 2, it is seen that, on a continuous (nonsampling) basis, a $K_v=0.97$ is predicted by the pseudo system for a 20-per-cent overshoot, compared to a K_v of unity for the actual system (as determined with an analogue computer). With reference to Figs. 3 and 4 for the continuous case, it is seen that the pseudo system predicts a $t_p=3.9$ as compared to a $t_p=4.2$ for the actual system and a $t_r=7.4$ as compared to a $t_r=8.2$ for the actual system. These results are within the 10 per cent accuracy expected.

EXAMPLE 2—OPEN-LOOP TRANSFER FUNCTION CONTAINING REAL AND COMPLEX ROOTS

The open-loop transfer function described by equation 5 is an example of the category of control systems containing a minor feedback loop.

$$G(s)=\frac{K_v}{s(s+1)\left(\frac{s^2}{4}+\frac{s}{2}+1\right)} \tag{5}$$

Equation 5 is seen to contain a dominant real root, and a significant pair of complex roots with a damping ratio $\zeta=0.5$. The real part of $j\omega G(j\omega)$ is plotted in Fig. 7 for this case, along with the pseudo system that resulted in the closest correlation between real-part plots. The pseudo system so chosen has an $n=4$ and a $\tau=0.375$ (obtained by matching crossover frequencies). Fig. 8 shows the correlation between step-function responses for the actual and the pseudo systems. The top graph corresponds to the case wherein the error is supplied continuously ($T=0$), and the middle graph shows the results when the error is sampled at 1-second intervals ($T=1.0$ second). These transient response curves were obtained by analogue computer studies.

The transfer function defined by equation 5 was also analyzed to consider the effect of a 0.5-second time delay for a sampling period of 1 second ($T=1.0$ second). The time delay considered here is of the form that arises in the case of a control system whose error is being supplied in sampled-data form from a digital computer link, assuming that the time delay between sampling the inputs and computing the error is the same for all the input variables involved. To select a pseudo system for this case, the time delay was incorporated with the continuous portion of the modified open-loop transfer function. The real part of $j\omega G(j\omega)\epsilon^{-0.5j\omega}$ was plotted in Fig. 7 and a new pseudo system was selected in the same fashion as before. The pseudo system for this case has an $n=4$ and a $\tau=0.5$ (by matching frequencies of crossover). The bottom graph in Fig. 8 shows the step-function responses of the actual and pseudo systems for this case, as obtained by analogue computer studies.

It should be noted that all the results plotted in Fig. 8 were obtained by adjusting the system K_v for a 20-per-cent peak overshoot in step-function response. The numerical values of K_v shown in Fig. 8 were derived from the $n=4$ curve in Fig. 2, using the values for τ as derived from the real-part plots. The response curves in Fig. 8 thus represent a comparison between the actual system response and the response predicted by the pseudo system.

As discussed in the following section, the predicted behavior of the system is specified by delineating specific features of the transient response such as rise time and response time. The curves in Fig. 8 serve to emphasize that the pseudo system response actually describes the system response accurately throughout the transient period.

Specification of Transient Response

The transient response of an underdamped control system can be specified in terms of the following three quantities:

1. Per-cent peak overshoot.
2. Time to peak of first overshoot (t_p), as measured from the instant that the error is first sampled.
3. Time to peak of second overshoot (t_r), as measured from the instant that the error is first sampled. (The use of this quantity as a measure of settling time presumes a knowledge of the relationship between the magnitudes of the first and second overshoots, as will be discussed.)

It should be noted that for the case wherein the error is supplied in sampled-data form, the response times t_p and t_r could be increased by a time delay t_d where $t_d<T$. This added time delay is the random delay between the time that the step-function input is applied and when it is first sampled and presented as an error to the system.

Fig. 9 shows a typical underdamped step-function response of a type-1 control system with the foregoing quantities specified. It will be shown later that, for the control systems under consideration, the magnitude of the second overshoot will always be 5±1 per cent if a peak overshoot of 20 per cent is selected as the criterion for stability.

The root loci of the closed-loop poles of the four pseudo systems are shown in Fig. 10, plotted as a function of dimensionless loop gain (K_w) in the dimensionless w plane. The indicated poles correspond to the loop gain required for a 20-per-cent peak overshoot in the continuous transient response. It can be seen from Fig. 10 that, on the basis of a 20-per-cent peak overshoot, the decrement factor, or real part, of the dominant pair of complex roots is small compared to any other decrement factors present in the closed-loop response. This is to say that the contribution to the response by the nondominant roots will be negligible after the time required to reach the peak of the first overshoot. It can be concluded that the response, after the time to the peak of the first overshoot, will be completely specified by the dominant pair of complex roots. This fact is born out further by observing that the damping ratio ζ of the dominant pair of complex roots for each of the pseudo systems is approximately 0.4, on the basis of a 20-per-cent peak overshoot in step-function response, as compared to 0.42 for a simple quadratic system. Since a 20-per-cent peak overshoot for a single pair of complex roots gives a 5-per-cent second overshoot, and the responses of the pseudo systems are dominated by a single pair of complex roots, the magnitude of the second overshoot will always be approximately 5 per cent on the basis of a 20-per-cent peak overshoot in step-function response. This observation can also be applied to any control system whose open-loop real part can be simulated by a pseudo system, since the real-part plot uniquely specifies the transient response.

To describe adequately the settling time when the error is supplied in sampled-data form, the curves in Fig. 4 show the relationship between the time to the peak of the second overshoot (t_r) and the sampling period T, for the four pseudo systems, as determined by analogue computer studies. It was observed that, on the basis of a 20-per-cent peak overshoot, the second overshoot was 5±1 per cent even when the error was supplied in sampled-data form. A justification of this observation appears in Appendix II pertaining to the case of extremely long sampling periods.

Conclusions

An approximate technique has been presented for determining the transient response of an important class of sampled-data control systems. The technique offers a radical simplification of analysis, and yields information as to the permissible velocity constant for a specified transient response, at any sampling rate of interest.

Data presented in the paper was derived by the use of an analogue computer. Over-all accuracies are in keeping with the approximate nature of the method, and have been shown to be reliable to within 10 per cent.

The content of the paper lends itself well to stability studies of systems employing a digital computer link, including time delay arising from a digital computation of error.

The authors feel that further investigations should be made to exploit fully the potentialities of the technique. One problem of interest, not treated in this paper, is an extension of the pseudo system technique to predict the effects of compensation networks as applied to sampled-data systems.

Appendix I

The impulse response of a system transfer function $H(s)$ can be related to the real part of $H(j\omega)$, subject to the following restrictions

1. $ReH(j\omega)=ReH(-j\omega)$, where Re signifies the real part of
2. $ImH(j\omega)=-ImH(-j\omega)$, where Im signifies the imaginary part of
3. $\lim_{s\to\infty} H(s)=0$
4. $H(s)$ has no poles in the right-half s-plane, or on the imaginary axis.

Restrictions 1 and 2 as stated are a modification of the original restriction as imposed by Floyd[8] which states that $H(s)$ must be written as the ratio of two rational polynomials in s with real and constant coefficients. This statement imposes an unnecessary restriction upon the system transfer function, such as when $H(s)$ includes a time delay within the major feedback loop or when the pseudo system is characterized by a noninteger value of n. Both of these cases will be discussed.

The impulse response of a system transfer function $H(s)$, that satisfies the above restrictions, can be uniquely related to the $ReH(j\omega)$ as follows[8]

$$h(t)=\frac{2}{\pi}\int_0^\infty [ReH(j\omega)\cos t\omega]d\omega \qquad (6)$$

where $h(t)$ is the impulse response of the system transfer function, $H(s)$.

Since a control system with time delay within the major feedback loop is physically realizable, and restrictions 1 and 2 are a necessary condition for physical realizability, equation 2 can be applied to transfer functions that fall within this category.

To demonstrate the applicability of equation 2 to a pseudo system with a noninteger value of n, it is only necessary to show that this case satisfies restrictions 1 and 2 since the other restrictions are clearly satisfied. The real part of the open-loop pseudo-system transfer function, with the integration removed, can be written in dimensionless form as follows, for $K_u=1$

$$Re[juG(ju)]=\frac{\cos n\tan^{-1}u}{(1+u^2)^{n/2}} \qquad (7)$$

And since

$$\frac{\cos n\tan^{-1}(-u)}{[1+(-u)^2]^{n/2}}=\frac{\cos n\tan^{-1}u}{(1+u^2)^{n/2}} \qquad (8)$$

for all n, the $Re[(ju)G(ju)]$ is an even function of frequency. Likewise

$$Im[juG(ju)]=\frac{\sin n\tan^{-1}u}{(1+u^2)^{n/2}} \qquad (9)$$

And since

$$-\frac{\sin n\tan^{-1}(-u)}{[1+(-u)^2]^{n/2}}=\frac{\sin n\tan^{-1}u}{(1+u^2)^{n/2}} \qquad (10)$$

for all n, the $Im[(ju)G(ju)]$ is an odd function of frequency, and equation 6 can be applied when n is not an integer.

Appendix II

In each of the Figs. 2, 3, and 4, defining the velocity constant and response characteristics of the various pseudo systems as a function of sampling period, broken-line curves have been plotted representing the behavior of the pure integrator (i.e., the pseudo system wherein $n=0$). These curves are of interest, since any open-loop transfer function receiving sampled error will behave like a pure integrator for sampling periods long compared to the time constants of the system. Hence, certain trends in the data, pertaining to the pseudo system response at long sampling periods, may be analyzed with reference to the pure integrator response.

In Fig. 3, it is seen that the curve for $n=1$ describing the time to the peak of the first overshoot, has an inflection at a dimensionless sampling period $T/\tau=6$. This may be explained by noting that for still larger values of T/τ, the curve seeks the pure integrator response. In the remaining curves in Fig. 3, for $n=2$, 3, and 4, no such inflection is observed for the values of T/τ plotted. It is inferred that inflections in the curves for $n>1$ will occur at larger values

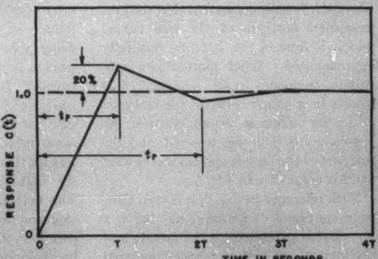


Fig. 11. Step-function response of the pure integrator case when the error is supplied in sampled-data form

of T/τ, since larger values of n impose more restraints on the initial value of the response.

With respect to Fig. 4 which relates t_r to sampling period, it can be shown that the magnitude of the second overshoot will approach 4 per cent for sampling periods very long compared to any time constants in the open loop transfer function. This is the case where all type-1 systems respond approximately as a pure integrator. For this case, with reference to Fig. 1

$$G(s)=K_v/s \tag{11}$$

where $K_v=1.2/T$ so that the peak overshoot in unit step-function response be 20 per cent. The error will be a superposition of step functions dependent on the value of error sampled at each sampling instant, described by

$$E_2(s)=\frac{A_0}{s}+\frac{A_1\epsilon^{-Ts}}{s}+\frac{A_2\epsilon^{-2Ts}}{s}+\ldots+\frac{A_n\epsilon^{-nTs}}{s} \tag{12}$$

where

$$A_0=1$$

$$A_n=C[(n-1)t]-C(nt) \text{ for } n\geq 1 \tag{13}$$

In equation 13, $C(nT)$ equals $C(t)$ evaluated at $t=nT$. Therefore,

$$C(s)=[E_2(s)]G(s)=\frac{1.2}{T}\left[\frac{A_0}{s^2}+\frac{A_1\epsilon^{-Ts}}{s^2}+\ldots+\frac{A_n\epsilon^{-nTs}}{s^2}\right] \tag{14}$$

Since we are interested in the value of $C(t)$ evaluated at $t=2T$ (the value of the second overshoot)

$$C(t)=\frac{1.2}{T}\,[t-1.2(t-T)] \tag{15}$$

for $T\leqslant t\leqslant 2T$. Consequently

$$C(2T)=\frac{1.2}{T}\,[2T-1.2T]=0.96$$

which indicates that the magnitude of the second overshoot is 4 per cent independent of the sampling period T.

Fig. 11 is a plot of the unit step-function response of any type-1 system, having a sampling period long enough so that it responds approximately as a pure integrator.

References

1. Sampled-Data Control Systems Studied Through Comparison of Sampling with Amplitude Modulation, **William K. Linvill.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1779–88.

2. The Analysis of Sampled Data Systems, **J. R. Ragazzini, L. A. Zadeh.** *AIEE Transactions*, vol. 71, pt. II, Nov. 1952, pp. 225–34.

3. Principles of Servomechanisms (book), **G. S. Brown, D. P. Campbell.** John Wiley & Sons, Inc., New York, N. Y., 1948, pp. 332–56.

4. Servomechanism and Regulating System Design (book), **H. Chestnut, R. W. Mayor.** John Wiley & Sons, Inc., New York, N. Y., vol. I, 1951, pp. 345–6.

◆

No Discussion

The Influence of Time Scale and Gain on Criteria for Servomechanism Performance

DUNSTAN GRAHAM
MEMBER AIEE

RICHARD C. LATHROP
ASSOCIATE MEMBER AIEE

THE mathematical synthesis problem in linear servomechanism design is to find the description of the system when the desired output response to a given input is specified. The desired output, however, is often defined as the best that can be obtained, where the word "best" has several connotations which include fastest transient response, minimum overshoot, and minimum tendency toward persistent oscillations. Under the assumption that there are no restrictions on the values that a sufficient number of system parameters may take, it is possible to apply suitable criteria and select the optimum response in a normalized time scale. The optimum response may then, theoretically at least, be reproduced, by adjustment of the system parameters, to an arbitrary time scale.

The selection of optimum responses in normalized time has been demonstrated in two previous articles,[1,2] for a variety of systems up to the eighth order. A number of possible criteria on which an optimum selection might be based were investigated in reference 1. Of these, the integral of time-multiplied absolute-value of error (ITAE) criterion was demonstrated to be preferable to others with respect to reliability, selectivity, and ready applicability. This criterion recognizes that in evaluating transient response the sign of a servomechanism error is not important and that an initial error must be accepted, but that the system should be penalized for any persistent error. Other feasible and widely used criteria are the solution-time criterion and the integral of squared error (ISE) criterion.

In practical servomechanism synthesis, however, more often than not the number or range of system parameters available to the designer for adjustment is inadequate, and the adjustment for optimum response cannot be independent of the gain adjustment which controls the time scale. Under these circumstances the implications of applying a unitary criterion are less obvious and the character of an optimum response is more obscure. It is, for example, widely known that an increase in the loop gain of a servomechanism usually tends to contract the time scale of the transient response and to decrease the relative stability. However, there is by no means any general agreement on the most desirable compromise between these two tendencies. Dr. Paul E. Pfeiffer has suggested, in a private communication, that the conclusions drawn in reference 1 about the merit of several criteria and the responses which they select as optima might be altered by evaluating the criteria in real time as a function of the open-loop parameters, including the gain. The effects on several criteria of changes in time scale and in the loop gain by itself, with particular reference to general concepts of the desired response of a servomechanism, are explored in this paper.

Symbols and Definitions

A_i, B_i=normalized transfer function coefficients
a, b, c, d=constants of the characteristic equation
$C(p)$=Laplace transform of the output (response)
e=error=input minus output
K=open-loop gain
$R(p)$=Laplace transform of the input
p=complex Laplace variable
P_i, Q_i=transfer function coefficients
$s=p/\omega_0$=normalized Laplace variable
t=time
ζ=damping ratio
τ=time constant
ω_0=natural angular frequency
$\text{ISE}=\int_0^\infty e^2dt$

Paper 54-197, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE North Eastern District Meeting, Schenectady, N. Y., May 5–7, 1954. Manuscript submitted February 4, 1954; made available for printing March 25, 1954.

Dunstan Graham and Richard C. Lathrop are with the All Weather Branch, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

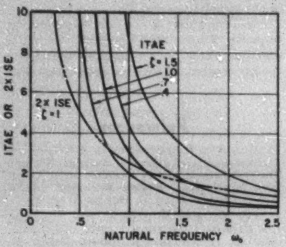


Fig. 1. Effect of relative time scale or system natural frequency upon the value of the ITAE and ISE criteria, second-order systems

$\text{ITAE}=\int_0^\infty t|e|dt$

Sol. Time = solution time, time required for error to reach and remain within only 5 per cent of its initial value following a specified disturbance

Time Scale

The time scale of the transient response of a servomechanism may be normalized by the substitution of a new independent variable $t'=\omega_0 t$ in the differential equation which describes the system. In terms of the general servomechanism system transfer function

$$\frac{C(p)}{R(p)}=\frac{P_m p^m+\ldots+P_2 p^2+P_1 p+P_0}{Q_n p^n+\ldots+Q_2 p^2+Q_1 p+Q_0} \tag{1}$$

$$\omega_0^n=Q_0/Q_n \tag{2}$$

Let

$$A_i=\frac{P_i}{\omega_0^{n-i}Q_n} \qquad i=0, 1, 2, \ldots m \tag{3}$$

$$B_i=\frac{Q_i}{\omega_0^{n-i}Q_n} \qquad i=1, 2, 3, \ldots n \tag{4}$$

then

$$\frac{C(p)}{R(p)}=\frac{A_m\omega_0^{n-m}p^m+\ldots+A_1\omega_0^{n-1}p+A_0\omega_0^n}{p^n+B_{n-1}\omega_0 p^{n-1}+\ldots+B_1\omega_0^{n-1}p+\omega_0^n} \tag{5}$$

Further, let

$$s=p/\omega_0 \tag{6}$$

then

$$\frac{C(s)}{R(s)}=\frac{A_m s^m+\ldots+A_1 s+A_0}{s^n+B_{n-1}s^{n-1}+\ldots+B_2 s^2+B_1 s+1} \tag{7}$$

The normalized transfer function equation 7, in which the coefficients of the first and last terms in the denominator are both unity, corresponds for a selection of values for A_i and B_i to a unique transient response (for a given input) to the scale of normalized time $\omega_0 t$. The time scale of the response may be expanded or contracted without altering other aspects of the system dynamics provided the parameters A_i and B_i are multiplied by suitable powers of the time scale factor ω_0 as indicated by the coefficients of the various powers of the complex variable p in equation 5. This has the same apparent effect as producing the responses experimentally and recording the results at fast and slow paper speeds.

Thus, for example, in the case of the linear second-order system with the transfer function

$$\frac{C(p)}{R(p)}=\frac{\omega_0^2}{p^2+2\zeta\omega_0 p+\omega_0^2} \tag{8}$$

corresponding to equation 5, or

$$\frac{C(s)}{R(s)}=\frac{1}{s^2+2\zeta s+1} \tag{9}$$

corresponding to equation 7, the form of the response depends only on the damping ratio ζ. The time scale of the response may be expanded or contracted by selecting a lower or a higher value for the natural frequency ω_0 in the unnormalized transfer function expression of equation 8, while the damping ratio ζ is fixed.

Qualitatively, except for considerations of servomechanism response to noise, there is no room to doubt that a contracted time scale or faster response always represents an improvement in performance. Quantitatively, there may be some difference of opinion about whether or not halving the time scale (doubling the natural frequency) improves the performance by exactly a factor of 2, no matter how far this process is carried. Raising the natural frequency of a servomechanism from 2 to 4 cycles per second, while maintaining the relative stability by other adjustments, may represent a considerable improvement, while raising the natural frequency of the same servomechanism from 4 to 8 or from 8 to 16 cycles per second might not be held to be as relatively worth while (especially with regard to the probable cost of such an undertaking). Such an estimate of value, however, is made with respect to the time scale of other phenomena, such as the input signal or the response of the controlled member or plant. It is likely to be more realistic than one arrived at by considering the servomechanism to exist in a vacuum.

Fig. 1 illustrates the variation of two integral criteria applied to the step-function responses of a linear second-order servomechanism, with the transfer function of equation 8, as the time scale of the responses is varied. The shape of the curve representing the ISE criterion for a damping ratio $\zeta=1$ is that of a hyperbola, and illustrates the species of criterion which halves its value each time the time scale is halved or the natural frequency is doubled. The solution-time criterion is also of this kind since its value, which has the dimension of time, is bound to be exactly inversely proportional to the time scale factor. On the other hand, the ITAE criterion demonstrates the evaluation of relative merit dis-

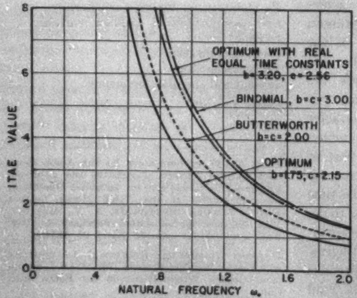


Fig. 2 (left). Effect of relative time scale or system natural frequency upon the value of the ITAE criterion, third-order systems

$$\frac{C}{R}=\frac{\omega_0^3}{p^3+b\omega_0 p^2+c\omega_0^2 p+\omega_0^3}$$

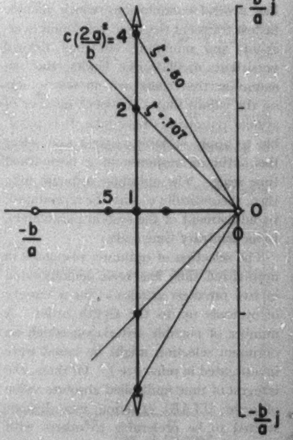


Fig. 3 (right). Locus of roots of the equation $ap^2+bp+c=0$ as c varies between zero and infinity

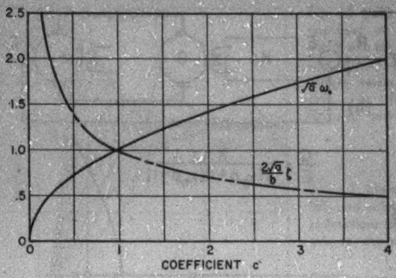
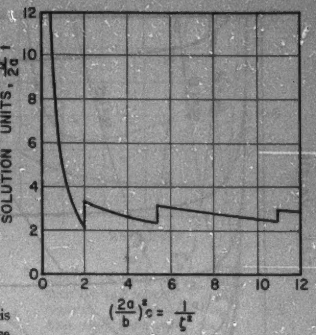


Fig. 4 (left). Dependence of the damping ratio ζ and natural frequency ω_0 upon the coefficient c for systems with the characteristic equation $ap^2+bp+c=0$

Fig. 6 (right). The solution-time criterion as a function of the coefficient c in systems with the transfer function $C/R = c/(ap^2+bp+c)$ for step-function inputs



cussed previously, which shows a decreasing proportional benefit as the natural frequency is continually increased.

When it is applied to third-order systems the ITAE criterion indicates very much the same things. The rank of several responses is unchanged as the time scales are contracted or expanded together and the criterion decreases in value for each response more and more slowly as the natural frequency is increased further and further. This is illustrated in Fig. 2 where the variation of the ITAE criterion is plotted as a function of natural frequency for several well-defined types of third-order system step-function responses.

The data presented in Figs. 1 and 2 may also be considered to represent, in a more general way, the relative merit of contracting or expanding the time scale, from a given starting point, if the abcissa is taken to be a frequency ratio.

Gain

The closed-loop natural frequency of a servomechanism is a function of the open-loop gain, and the gain is often one of very few, sometimes the only, system parameters which are readily accessible to the designer for adjustment. In any system except one of the first order this would mean that a given form of response could not be produced to an arbitrary time scale, and that some compromise between a fast response and decreased relative stability is indicated. The analysis of the effects of increasing loop gain is extensively discussed in the literature of closed-loop control systems. The root-locus plot and the Nyquist diagram are particularly amenable to the determination of the effects of gain variation as such. Unfortunately, except for stability and the rather arbitrary but widely accepted figure for maguitude ratio $M_p=1.3$, there are no standards for an optimum synthesis against which the results of selecting a particular value of the open-loop gain may be compared. It is interesting to contrast what the criteria for transient performance show with the performance of servomechanism systems, as a function of gain, as indicated by other methods.

Fig. 3 illustrates the locus of roots of the characteristic equation

$$ap^2+bp+c=0 \qquad (10)$$

as the coefficient c varies from zero toward infinity. This characteristic equation is that of the simple second-order system illustrated in block-diagram form in the inset of Fig. 7. The closed-loop transfer function of this system is

$$\frac{C(p)}{R(p)}=\frac{K}{\tau p^2+p+K}=\frac{c}{ap^2+bp+c} \qquad (11)$$

or

$$\frac{C(p)}{R(p)}=\frac{K/\tau}{p^2+\frac{1}{\tau}p+K/\tau}=\frac{\omega_0^2}{p^2+2\zeta\omega_0 p+\omega_0^2} \qquad (12)$$

Fig. 3 shows that no matter how much the gain is increased the system never does become unstable, although the damping ratio decreases as the gain and natural frequency are increased. This type of behavior is peculiar to second-order systems.

Fig. 4 illustrates the actual variation of the damping ratio and natural frequency as functions of the gain parameter c with other system constants held fixed. The results of varying the gain to obtain a higher natural frequency ω_0 and a lower damping ratio ζ on the one hand, or a lower natural frequency and a higher damping ratio on the other hand, do not easily lead to conclusive statements concerning relative merit. The $M_p=1.3$ criterion of frequency-response synthesis leads to selecting the rather low damping ratio $\zeta=0.42$ for second-order systems.

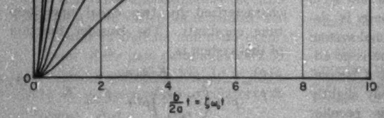
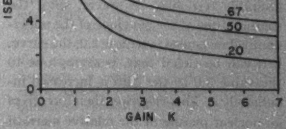


Fig. 5 (left). Step-function responses of a system with the transfer function $C/R = c/(ap^2+bp+c)$ for various values of the coefficient c

Fig. 7 (right). Variation of ISE criterion value with open-loop gain and time constant, simple second-order systems



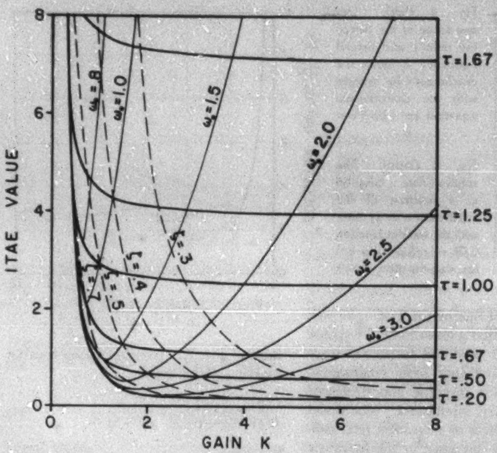


Fig. 8. Variation of the ITAE criterion value with open-loop gain and time constant, simple second-order systems

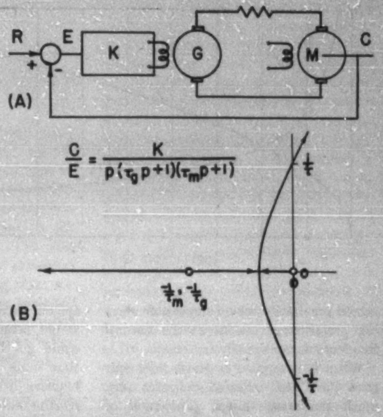


Fig. 9. A—Schematic diagram of a third-order servomechanism. B—Locus of roots of the system of A as the open-loop gain is increased, for the special case $\tau_g = \tau_m$

The step-function responses of simple second-order systems with the transfer function of equation 11 are illustrated for various values of the gain parameter in Fig. 5. The curves are universal for systems of this type because both the time and the gain parameter have been nondimensionalized. It may be seen that the first overshoots and subsequent undershoots and overshoots of the underdamped responses are always contained within a very close approximation to an exponential envelope. The time constant of the approximate envelope[3] is $2a/b = 1/\zeta\omega_0$. This indicates that even in second-order systems, where the possibility of making the servomechanism unstable does not exist, there is no virtue, as far as increasing the speed of response is concerned, in raising the gain beyond a value which gives an only slightly underdamped response. Fig. 6 illustrates this point very neatly. It shows a plot of the solution-time criterion applied to the responses of Fig. 5. The nondimensional solution time is shown as a function of the nondimensional gain parameter. This criterion selects the value of gain corresponding to a damping ratio $\zeta = 0.7$ as optimum, and this selection agrees with widely accepted concepts of a desirable response. However, since the solution time of the envelope of the underdamped responses is $bt/2a = 3.0$, the curve, if it can be called that, is asymptotic to the value 3.0 as the gain is increased indefinitely. This criterion therefore gives an unclear picture of the relative merit of simple second-order systems as the gain is increased beyond the value which yields a damping ratio $\zeta = 0.7$.

Unfortunately neither the ISE, $\int_0^\infty e^2 dt$, nor the ITAE, $\int_0^\infty t|e| dt$, criteria provide a better picture. These criteria have been evaluated in real time for unit step functions of input to the system of equation 11. The results are presented in Figs. 7 and 8 as functions of the open-loop gain K with the time constant τ as a parameter. The two criterion surfaces, as functions of K and τ, show some topological similarities. At a constant value of time constant both criteria decrease in value as the gain K is increased, and at a constant value of the gain both criteria decrease as the time constant is decreased. The ITAE criterion strongly emphasizes the desirability of a low value for τ and the onset of diminishing returns as the gain is increased indefinitely. The fact is, however, that the knee in the curves occurs at widely differing nondimensional damping ratios. This may be observed in connection with the contours of constant damping ratio ζ, and natural frequency ω_0, which are superposed on Fig. 8. If the dimensional system damping is inherently low and the gain is the only possible adjustment, it would seem that an underdamped response is demanded, while if the dimensional system damping is high (τ small) there is no advantage in increasing the gain K beyond a value which produces an only slightly underdamped response. This peculiar conclusion is not necessarily wrong, but it would have to be tempered in practice by consideration of power requirements and the intended application of the servomechanism.

For variations in the second-order system parameters which might be practically available, the ISE and ITAE criteria correctly indicate the direction of improved performance except for a lack of discrimination against the oscillatory responses. The latter failing would probably be common to all servomechanism performance criteria which are a function of error since, as has already been noted, the responses of simple second-order systems tend toward being uniformly contained in an exponential envelope whose shape is not a function of the increasing gain.

In higher order systems, however, increasing the gain eventually makes the servomechanism dynamically unstable, and criteria based on the error have no undue difficulty in effecting a choice short of that undesirable condition. While it is relatively difficult to make general statements regarding higher order systems, an illustrative example of a third-order system will serve to show the effect on the response and on the criteria of variations in the gain. Fig. 9(A) illustrates a third-order system which is characterized by two equal open-loop time constants. The transfer function of this system is

$$\frac{C(p)}{R(p)} = \frac{K/\tau_m\tau_g}{p^3 + \left(\frac{\tau_m + \tau_g}{\tau_m\tau_g}\right)p^2 + \frac{p}{\tau_m\tau_g} + \frac{K}{\tau_m\tau_g}} \quad (13)$$

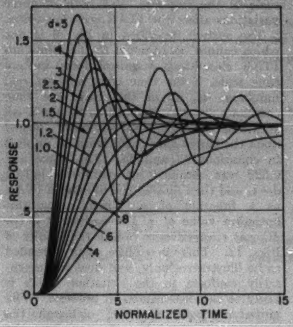


Fig. 10 (left). Step-function responses of the system of Fig. 9 with the transfer function $C/R=d/(p^3+bp^2+cp+d)$ for various values of open-loop gain. In this figure $d=K/(\tau_m\tau_g)$ and $\tau_m=\tau_g=0.625$

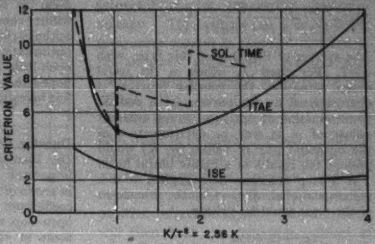


Fig. 11 (right). The ITAE, ISE, and solution-time criteria as functions of the open-loop gain for the system of Fig. 9

or

$$\frac{C(p)}{R(p)}=\frac{d}{p^3+bp^2+cp+d} \tag{14}$$

For a gain parameter $d=1.0$, the optimum values of b and c, as selected by the ITAE criterion and subject to the condition that $b^2=4c$, are $b=3.20$ and $c=2.56$. These conditions and values correspond to $\tau_m=\tau_g=0.625$. Fig. 9(B) shows the locus of roots, starting from $-1/\tau_m=-1/\tau_g=-1.60$ and the origin, as a function of gain. The tendency of the system to become unstable is indicated by the two branches which cross the imaginary axis into the right half plane.

Fig. 10 shows the corresponding unit step-function responses and clearly illustrates the initial rapid improvement in speed of response as the gain is increased and the subsequent tendency for the system to become unstable as the gain is increased still further.

Fig. 11 shows three curves which correspond to the variation of three criteria when they are applied to the responses of Fig. 10. Both the solution-time criterion and the ITAE criterion are selective and reliable in the sense that their rather sharp minimum values occur at a value of gain corresponding to a transient response which shows good qualities of speed of response and minimum overshoot. Note, however, that the discontinuities in the solution-time criterion curve give a false impression of the merit of systems with very slightly different characteristics. This feature tends to make the criterion unamenable to rational optimizing procedures. The ISE criterion is unselective in that it has approximately the same value for a wide variety of responses, and is unreliable in that its minimum occurs at a value of the gain parameter $d=2.50$, which may be seen from Fig. 10 to correspond to an excessively oscillatory response.

Conclusion

The suggestion that unitary figures of merit for servomechanism performance might tell a different story if they were evaluated in terms of real time rather than of normalized time has not been borne out by this investigation. Except for the fact, peculiar to second-order systems as the gain parameter is increased, that none of the criteria discussed are capable of sharply discriminating against the more oscillatory second-order responses, the variation of the criteria with time scale and gain changes shows no very remarkable features. The solution-time criterion retains its undesirable discontinuities. The ISE criterion still shows a lack of selectivity and a preference for underdamped responses. The ITAE criterion, however, retains the advantages of selectivity and reliability claimed for it in previous reports, and it would seem to commend itself for the evaluation of transient responses and the choice of optima in real and normalized time.

References

1. THE SYNTHESIS OF "OPTIMUM" TRANSIENT RESPONSE: CRITERIA AND STANDARD FORMS, **Dunstan Graham, R. C. Lathrop.** *AIEE Transactions,* vol. 72, pt. II, Nov. 1953, pp. 273–88.

2. THE TRANSIENT PERFORMANCE OF SERVOMECHANISMS WITH DERIVATIVE AND INTEGRAL CONTROL, **Richard C. Lathrop, Dunstan Graham.** *AIEE Transactions,* vol. 73, pt. II, Mar. 1954, pp. 10–17.

3. SERVOMECHANISM ANALYSIS (book), **G. J. Thaler, R. G. Brown.** McGraw-Hill Book Company, Inc., New York, N. Y., 1953.

4. AUTOMATIC FEEDBACK CONTROL (book), **W. R. Ahrendt, J. F. Taplin.** McGraw-Hill Book Company, Inc., New York, N. Y., 1951.

◆

Discussion

Paul E. Pfeiffer (The Rice Institute, Houston, Tex.): In presenting the series of studies on the properties of the ITAE criterion, the authors are making a valuable and stimulating contribution to the profession. They have shown in many important cases that the ITAE criterion is superior to other proposed criteria on the basis both of selectivity and of the kind of response it selects as the "optimum." Also, they have made use of the idea of "standard forms" for which they have presented many valuable data.

In regard to their use of these standard forms, certain questions arise in my mind. Their process of optimizing involves fixing the quantity ω_0 which is related to the closed-loop natural frequencies. In many situations it is desirable to specify closed-loop behavior, and then determine the system which will meet the requirements. On the other hand, there are many situations in which some open-loop time constant or natural frequency is fixed and various loop parameters are adjusted. One might be inclined to suppose that adjustment on either basis would lead to the same results. However, as the authors show, this is not necessarily the case. For example, the adjustment of the gain K of the second-order system of Fig. 7 does not lead to a distinct optimum set of parameters. On the other hand, the results contained in Fig. 6 of their first paper (reference 1 of the paper) do indicate such an optimum. A second example is the system described in Figs. 9 through 11 of the present paper. If the transfer function is normalized for the optimum value of $d=1.3$, the result yields coefficients (in terms of the normal forms of reference 1) $b=2.93$ and $c=2.15$, with $\omega_0=1.09$. Entry of these values into Fig. 12, reference 1, indicates the difficulty.

The discrepancy between the two approaches naturally raises the following questions: Which method of optimization leads to the lower ITAE value? Which leads to the better optimum response? The very fine results shown in Figs. 10 and 11 indicate that the answer to both questions is likely to be the same. There does not seem to be sufficient data to answer either question, however.

It seems to me that a table of standard forms based on the adjustment of open-loop parameters would be of considerable interest. For one thing, such standard forms would be more readily translated into system parameters. And it is conceivable, though not clear at present, that such standard forms would actually be superior to those determined on the closed-loop basis. A few well-chosen cases should shed considerable light on this matter.

The following example will indicate the kind of program that might be carried out. Consider a simple third-order system characterized by two time constants. The open loop transfer function for such a system is

$$G(s)=\frac{C(s)}{E(s)}=\frac{K}{s(Ts+1)(\alpha Ts+1)}$$

The system of Fig. 9 is of this type, with $\alpha=1$. For fixed T, which is the usual design situation, the transfer function is characterized by the parameters α and K. Determination of optimum K for different values of α as well as determination of the optimum pair of values (α, K) would present the results in a highly useful form.

If a step response were applied to the system, so that $R(s)=1/s$, the transform of the error could be written

$$\epsilon(w)=T\frac{\alpha w^2+(\alpha+1)w+1}{\alpha w^3+(\alpha+1)w^2+w+KT}=T\epsilon_0(w)$$

where $w=Ts$. Then if $\tau=t/T$, $e(t)=\epsilon(\tau)=T\epsilon_0(\tau)$, where $\epsilon_0(w)=L[\epsilon_0(\tau)]$.

In this case

$$I=\int_0^\infty t|e(t)|dt=T^3\int_0^\infty \tau|\epsilon_0(\tau)|d\tau=T^3I_0$$

The function $\epsilon_0(\tau)$ and hence I_0 varies with α and KT. Under the condition of fixed T, the optimum of α and K are chosen by minimization of I_0. Clearly this approach extends readily to more general cases.

It seems important to distinguish clearly between the use of the ITAE as a criterion of goodness and the use of the standard forms. The authors have presented ample evidence to justify confidence in the ITAE figure of merit as a guide in many common applications. The question that does not seem to be settled is whether standard forms should be based on closed-loop optimization with ω_0 fixed or on open-loop optimization with some time constant or open-loop natural frequency fixed. It is hoped that the authors will continue their fine investigation of the ITAE criterion.

Dunstan Graham and **Richard C. Lathrop:** Dr. Pfeiffer has correctly emphasized the important difference between optimizing in terms of the closed-loop transfer function coefficients and optimizing in terms of the open-loop system parameters. Our treatment of the first case, as developed in references 1 and 2 of the paper, has been based upon a normalization technique which involves separation of the dynamic properties of a response from the time scale at which the response occurs. In general, if a sufficient number of system parameters are available for adjustment through an arbitrary and perhaps complex range, it is then theoretically possible to adjust the closed-loop system to match a particular standard form, and to any desired time scale. If the number or range of parameter adjustments is restricted, however, the time scale may be similarly restricted, or it may be impossible to achieve the desired standard form to any time scale.

In the second case, one finds that adjustment of any one open-loop parameter results, in general, in a simultaneous variation of the dynamic properties and the time scale of the response, and both of these variations affect the ITAE criterion value.

The system described in Fig. 9 is one which cannot be adjusted to the minimum ITAE closed-loop standard form in any time scale, if the time constants are real numbers. If, as in this case, the time constants are also assumed to be equal, an additional restraint is obviously placed upon the closed-loop coefficients, as implied in equations 13 and 14. The value $\tau=0.625$ was obtained by arbitrarily letting $d=1$, and then finding the minimum ITAE as a function of τ. Then, using this optimum value of τ, the coefficient d (or K) was readjusted to obtain the results of Fig. 11. This procedure was intended to be illustrative only, and does not necessarily conform to the situation which is likely to be encountered in practice. The important point is that, although the system can never be adjusted into the optimum closed-loop standard form, the ITAE can still be used to obtain the best possible adjustment with the given restraints.

The normalization technique suggested by Dr. Pfeiffer for optimizing the simple third-order system in terms of the open-loop parameters K and α shows promise. We intend to evaluate this case, and also the fourth-order case. Unfortunately, it appears that these data do not lend themselves to as orderly a presentation as did the polynomial standard forms of the closed-loop coefficients.

Static Excitation Control for Diesel-Electric Locomotives

S. W. McELHENNY
ASSOCIATE MEMBER AIEE

R. M. SMITH
ASSOCIATE MEMBER AIEE

Synopsis: Recent trends in industrial control equipment have been toward the use of static devices. The objective has been to secure greater reliability, lower maintenance costs, and more accurate control. Absolute reliability is a prime requisite for railroad locomotive control. Substitution of easily serviced packages of static equipment for rotating machines and dynamic regulators of all types is one advance toward this ultimate of absolute reliability.

THE static excitation system here described has been developed for the main propulsion circuit of Diesel-electric road locomotives. This excitation control must impose three limits upon the generator-power plant combination:

1. A variable, controlled generator armature current limit.
2. A maximum Diesel-engine horsepower (hp) limit.
3. A maximum generator ceiling voltage limit.

Armature Current Limit

With the locomotive at standstill, the series traction motors present a low impedance to the generator and very little generator field current is required to produce high generator armature current. Under these conditions it is necessary to control generator armature current for two reasons: to limit starting traction motor torque, and to protect the generator and traction motors.

Too high a traction motor torque, the result of high motor current, can produce two undesirable results:

1. With high adhesion between the locomotive driving wheels and the rail, a drawbar fracture might well result from rapidly applied uncontrolled high motor torque.
2. With low adhesion, slipping driving wheels result.

Neither of these two conditions will successfully start a train.

Horsepower Limit

As the locomotive accelerates, the traction motors present an increasing load impedance to the generator. To maintain a constant generator armature current limit, it is therefore necessary to increase generator terminal voltage as the locomotive speed increases. This increasing volt-ampere product represents an increasing hp demand upon the Diesel engine, which, to maintain constant Diesel engine speed, necessitates an increase of fuel. Since there is a limit to the amount of fuel which can safely be fed to the engine, the electrical hp demand upon the engine may become greater than it can mechanically produce. Therefore, some control device must operate to match the electrical demands to the maximum hp which the engine is capable of delivering. If this is not done, the engine will stall down to some lesser speed. Since engine output varies directly with speed and generator output drops as the square of the speed, a balance will soon be reached at some

---

Paper 54-190, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE North Eastern District Meeting, Schenectady, N. Y., May 5-7, 1954. Manuscript submitted October 21, 1953; made available for printing March 1, 1954.

S. W. McElhenny and R. M. Smith are with the General Electric Company, Erie, Pa.

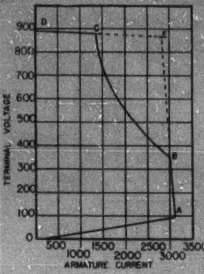


Fig. 1 (left). Generator characteristic for one throttle position

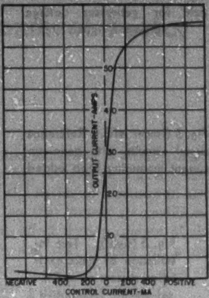


Fig. 3 (right). Amplistat transfer characteristics

hp level lower than the full-speed rating of the engine. This type of engine-stalling control is used extensively on small switching locomotives where maximum hp utilization is not important.

Ceiling Voltage Limit

The third limit imposed on the system is necessary for two reasons: To limit maximum generator field heating, and to limit maximum generator volatge. This is primarily for the protection of the electric equipment and does not enter directly into the control of the locomotive.

Fig. 1 plots generator terminal volts versus generator armature current for one throttle position, illustrating the foregoing. Line O*A* is the traction motor resistance line at standstill. Line *AB* shows the increase of terminal volts necessary to hold armature current limit as the locomotive accelerates. Line *BC* represents the maximum constant hp that the engine can safely deliver for this particular throttle position. To maintain constant hp it is necessary to hold the locus of current versus voltage to this line as the locomotive speed increases. Line *CD* represents the ceiling voltage portion of the characteristic. Very little locomotive acceleration will result, even with no train, after the generator reaches ceiling voltage limit, since the power delivered to the traction motors falls off very rapidly from *C* to *D*. At *C* it is necessary to shunt motor fields or change motor connections to increase speed further.

System Operation

The static excitation system consists of an alternator-rectifier combination with an amplistat (magnetic amplifier) which controls the output of the alternator to produce the limits just described. Fig. 2 schematically shows the basic system. The system is a closed loop feedback scheme which can be made to hold preset limits to a high degree of accuracy.

Fig. 3 shows the transfer characteristic of the amplisat. As will be noted, the amplistat exhibits a very high amplification over a portion of its characteristic, which makes it ideal for this type of control system. Positive control signals are to the right of the zero axis and negative control signals are to the left.

The positive reference signal for the system is provided by rectified, regulated alternator voltage.

Feed-back control signals are received from two different sources. The first, as shown in Fig. 4, is a series-connected saturable reactor measuring generator voltage. The output of this voltage-measuring reactor is rectified and fed to the amplistat negatively. When a large positive reference signal is fed to the amplistat, the amplistat will saturate, thereby increasing generator field current and, in turn, generator voltage. This increase of generator voltage will continue until the negative, rectified voltage-measuring reactor output balances the positive reference signal. The amplistat will then hold the generator voltage at this value until a change of control ampere turns on the amplistat occurs. Thus a regulated generator ceiling voltage limit has been achieved.

The second negative signal source is another series-connected saturable reactor physically located so that the main generator cables pass directly through it. The reactor, therefore, has an output which is proportional to generator current. The output of this current-measuring reactor is rectified and fed to the amplistat, as shown in Fig. 5.

The two negative signal sources are so connected in series that only one of them will be in control at any time. The two controlling elements will then produce a generator volt-ampere characteristic like *DCEBA* of Fig. 1.

To demonstrate this, assume that the generator armature circuit is open. The current-measuring reactor output will therefore be zero. The generator voltage will then build up to its limit when the proper reference signal is applied to the amplistat, producing voltage *OD* in Fig. 1. If the generator armature circuit is then closed by inserting a high-impedance load, we may assume that sufficient current will flow to cause the generator to operate at *C*. The current-measuring reactor will then increase its output from zero to some finite value. Let us assume this to be less than the output of the voltage-measuring reactor. Since the voltage-measuring reactor rectifier and the current-measuring reactor rectifier are in series, and the current-measuring reactor output is less than the voltage-measuring reactor output, only the voltage-measuring reactor output will appear in control winding 2 of Fig. 5. Generator terminal voltage will therefore attempt to remain constant at the value it assumed during the time that the generator armature circuit was open. Actually, there will be a slight drop in generator terminal volts from *D* to *C*, because an error signal is required to increase field current as the generator is loaded.

Now, assume that generator current at *E* will produce a current-measuring re-

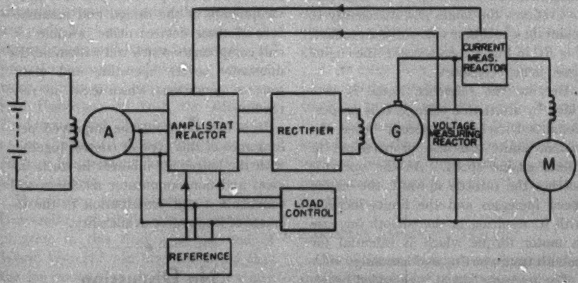


Fig. 2. Basic components of static excitation system

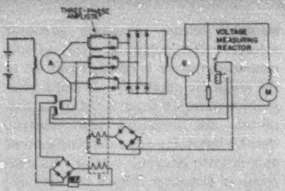


Fig. 4 (left). Feed-back circuit for generator voltage reference signals

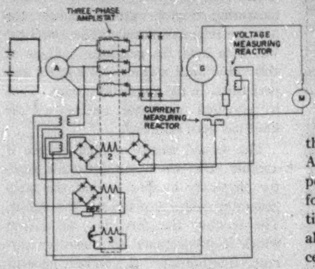


Fig. 5 (lower left). Reference signal circuits with generator current measuring loop added

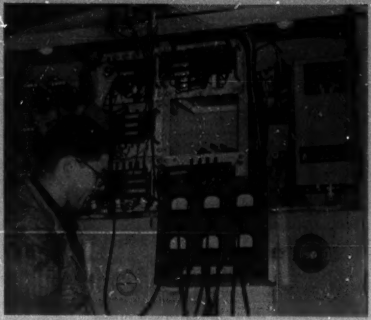


Fig. 6 (right). Test installation of static excitation system on a Diesel-electric locomotive

actor output equal to that of the voltage-measuring reactor. As the generator armature current exceeds the value of E, the current-measuring reactor output exceeds the voltage-measuring reactor output. Thus, at this point, the current-measuring reactor takes control; and its output, rather than that of the voltage-measuring reactor, appears in control winding 2. Since the current-measuring reactor output is now greater than the voltage-measuring reactor output, a negative increase in net control ampere turns to the amplistat is the result. This of course will decrease the amplistat output and cause a reduction in generator field current with a consequent reduction in generator terminal voltage. As the traction motors slow down they decrease load impedance to the generator, and the generator armature current rises if generator excitation is held constant. However, as it does so, the current-measuring reactor output rises, causing a reduction in generator terminal voltage. With a high-amplification system such as this, the armature current limit portion of the characteristic, line AE in Fig. 1, can be made to be nearly vertical since a very small error signal causes a large change in output.

A third limit on the system, which will be called load control, is necessary to convert the dotted characteristic $DCEBA$ of Fig. 1 into the solid line $DCBA$.

As previously mentioned, if the excitation is increased with the locomotive at standstill, the generator current and voltage will proceed along OA in Fig. 1 to A, assuming the armature current limit for this setting to be the value shown at A. Also, assuming conditions are such that point A represents sufficient tractive effort (motor torque) to start the locomotive, current and voltage will be held along the line AB as the locomotive accelerates. Because the engine governor attempts to keep a constant engine speed, it will increase fuel to the engine as voltage and hp build up along line AB. At B, a mechanical stop prevents the governor from feeding more fuel to the engine. After this point is reached, the electrical load progresses slightly along a projection of AB, which means more electrical demand than the engine can mechanically supply. This causes a governor-operated rheostat to produce a signal fed into the main amplistats. This is a negative signal and has the effect of reducing the reference signal. This, in turn, will reduce both the generator voltage limit and the armature current limit and thus will limit the electrical hp that the system is able to produce. For proper system operation, the load control must have sufficient range to be able to reduce the electrical output to a value lower than the maximum available engine hp.

As the load control signal increases, engine loading is reduced with no error in engine speed. Thus, as the locomotive accelerates above B, Fig. 1, the load control reduces the limits just sufficiently to maintain a constant volt-ampere product, line BC in Fig. 1. As a result, the engine speed is held constant.

Because the reference signal is provided by alternator voltage, it is proportional to Diesel engine speed; hence, the system limits are also proportional to Diesel engine speed. As the operator notches the throttle upward, the engine speed increases and the limits increase with it, resulting in the smooth increase in motor torque which is essential for smooth train-starting and handling.

The reference signal is provided by alternator voltage. A saturating reactor is added to compensate for alternator voltage change. This also allows the alternator to be used to provide a speed signal to the electrohydraulic engine speed governor.

Conclusion

This basic measuring system is flexible and can be modified to provide such features as an inherent electrical constant hp signal independent of load control action, controlled dynamic breaking, and constant track speed control for hump yard operation. The reliability is greatly enhanced by the relative absence of relay contacts and the substitution of an enclosed alternator for a d-c exciter.

Static excitation control has been successfully tested and is being applied to gas-turbine-electric and Diesel-electric locomotives. It eliminates the use of d-c commutator machines in the traction excitation circuits. The alternators which replace them promise to require much less maintenance.

The static devices used in the excitation circuits are not new. All of them—transformers, reactors, rectifiers, capacitors, and resistors—are being used in other types of successful and reliable circuits. Past operating experience plus recent developments in the design and manufacture of these devices make possible circuit components which will withstand the unusually severe operating and maintenance conditions which exist in railroading.

It is believed that the improved performance of this type of control together with the longer life inherent in static devices and noncommutator machines will constitute a real contribution to the increase of locomotive availability.

◆

No Discussion

Air Ionization as an Environment Factor

J. C. BECKETT
MEMBER AIEE

IN VIEW of the findings of research, an effort has been made to control ionization of air in the design of electric space-heating and air-conditioning equipment. Extensive studies show that ionization of the air has a physiological effect on people breathing the air, and it is reasoned that air-conditioning should not inadvertently change any property of fresh outdoor air if a change can reasonably be avoided.

Until additional biological experiments can show how to use ion control for relief of specific diseases or ailments, it is necessary in using the information learned thus far to be content with making indoor air like outdoor air at its best, before attempting to improve on the electrical properties of fresh air.

Interest in this work originated from the knowledge that electric heaters produce ions. With well-established proof that ionization is an environment factor affecting well-being and comfort, the next step was to apply this knowledge in the design of an ion-controlled electric heater.

When we speak of air-conditioning, we usually think in terms of circulating air combined with heating or cooling, filtering, and humidifying. These are the basic requirements for comfort in control of the atmosphere in which we live. In spite of the increased interest in research activity which has taken place since World War II, there still remains the question of the importance of air ionization as an environment factor.

Ions and Health

This subject of ionized air as affecting our health and comfort has been discussed for more than 50 years. It received considerable attention during the late 1920's and early 30's[1] and is now being actively studied again with the aid of greatly improved measuring techniques. Research work is diligently being pursued at Stanford University in both the Electrical Engineering and Chemistry Departments at the University of San Francisco in the Biology Department, and at the University of Arizona.

Hutchinson, mechanical engineer at the University of California, summarized progress in this field near the end of World War II.[2] In brief, he noted that the ion content of air varies with many factors and there is some evidence for believing that ionization is either partly responsible for, or in some indirect way related to, the feeling of freshness in air. He noted, too, that present air-conditioning methods affect the ion content of air appreciably and that heating engineers must understand how to control ions as ions may affect health and comfort.

Air-conditioning engineers are concerned with the methods of removing pollutants from the air with due regard to particle size. In addition to particulate matter, however, there is the influence of electrical charges on particles and molecules of gas making up the atmosphere. Uncharged particles below 1/10 micron in size behave differently from charged particles. What of these, and what of ions themselves? The charged particles and molecule complexes are commonly referred to as ions and may be classified into ion groups.

Ion Groups

There are three general groups of atmospheric ions: 1. Small ions are molecular complexes of a dozen or so gas molecules associated with one electrically charged molecule; 2. Intermediate ions generally exist in quantity near cities and areas of air pollution. They are submicroscopic particles believed to be formed where sulphuric acid particles from smokes are present or from hydrated ions;[3] 3. Large ions consist of clusters of small ions surrounding condensation nuclei or single charged particles of smoke. They are most common in urban areas or indoors where contaminates are present. Large ions are not particularly important as distinguished from uncharged particles because these can be filtered by mechanical means and the relative effect of one or two electrical charges on a particle of such great mass is negligible.

At present little is known about the intermediate ion because they are too small to be seen and difficult to isolate from the small ion group. It should be kept in mind that, though ions have been grouped in these three size ranges for convenience, there is no clear line of demarkation.

Particles Affecting Health

One of the most important physical factors in the causation of pulmonary disease is the size of particles, since only those below a certain size can reach the finer passages and air sacs of the lungs and start changes leading to disease.[4] Some of the particles that eventually reach the ultimate pulmonary tissue are deposited. It is known that this depositation within the lungs is selective and only particles below 5 microns penetrate to the lungs.[5] It has been observed that particles 1 micron in size most readily reach the lungs. Thirty-five per cent of the 0.1-micron sized particles are known to be retained. (These, if charged, would be classified as large ions.) Of particles below 0.1 micron about 25 per cent are retained. It is not safe to say that particles below 0.1 micron are biologically unimportant, as many engineers are prone to do. The behavior of particles of this small size (when electrically charged) is known to be different from their behavior when electrically neutral, owing to their small mass in proportion to the force of the electrical charge on them.

Ions Affecting Health

It is well known that particles deposited in the lungs affect health. Another question, still unsolved, is what effect, if any, does an electrical space charge in a room have on our feeling of well-being? To answer this question we must know something about the number of ions of each size found in free air, in closed rooms with and without air-conditioning, and about the magnitude of space charge that can be formed in a given room.

The Biological Aspects of Ions

Historically, considerable effort has been made to evaluate clinically the biological and physiological significance of atmospheric ions. Though these tests appeared to show a trend or pattern that was consistent, the results never withstood a careful statistical evaluation nor could the reports be duplicated at will. The principal difficulty appears to have been in trying to establish the facts by clinical methods in which control of the patients was inadequate to rule out other variables.

Biological studies with animals appear now to be the only practical solution for establishing the facts regarding the influence of ions on basic body functions. Once this has been accomplished, a carefully controlled clinical test can be run

---

Paper 54-225, recommended by the AIEE Domestic and Commercial Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing April 14, 1954.

J. C. Beckett is with the Wesix Electric Heater Company, San Francisco, Calif.

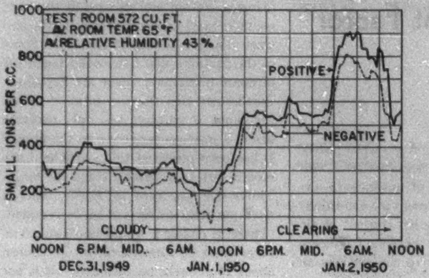


Fig. 1 (above). Typical small-ion density curves for room ventilated through open window, Pacific Heights, San Francisco

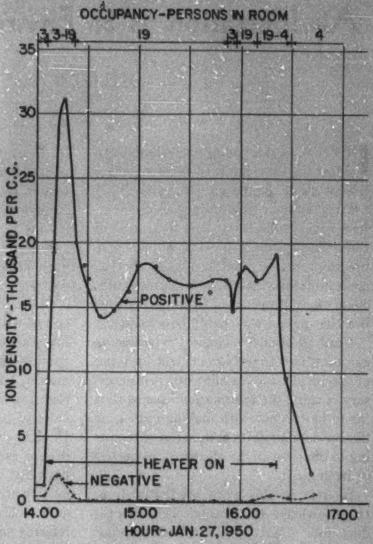


Fig. (right) 2. Room ion density characteristic with industrial-type 9KW electric fan heater

on some simple indicator such as the mucous membrane of the nose, throat, and trachea.

Early investigators finding it difficult to create relatively high ion densities free from ozone and oxides of nitrogen resorted to the use of MgO particles to reduce the mobilities of the ions and thereby permit higher concentrations. It was assumed the MgO was nontoxic and the particle size too small to be biologically harmful. It appears now that both assumptions were wrong to an extent that forces us to discard the results thus obtained.

If ions in the atmosphere are indeed important to consider in the design of air-conditioning equipment, it follows that the ions used for test, large or small, positive or negative, should be in quantities that might normally be found while working with such equipment.

Experiments with Animals

Biological experiments are still going on, but certain facts have been established:

1. Atmospheric ions at the relatively low densities that occur in free air do have a biological effect. This effect generally is so subtle on normal, healthy animals that it is necessary to use extremely sensitive indicators in order to establish the changes.
2. A sudden reversal in polarity of ionization momentarily disturbs the normal development and growth of animals.
3. Fundamental changes are mediated through the adrenal glands indicating possible therapeutic usefulness of ionized air.
4. Positive ions produce the most striking changes which are shown to be statistically significant.
5. Length of exposure is an important factor as animals adapt themselves to a given environment.

Ionization of the Air

When we speak of ionization of air as an environment factor we refer to the natural ion density as found in free air near the earth's surface the world over. This density varies from less than 200 ions per cubic centimeter (cc) to over 1,000 ions per cc. The positive small ions generally exceed the negative small ions by a slight amount in the ratio of 5 to 4.[6] This difference in number of ions results in a space charge in which the air of the room may be described as being at a potential different from that of the walls of the room (Fig. 1).

The small-ion density is observed to decrease at times of high relative humidity or air pollution but this does not account for all of the variations observed. In a series of tests conducted by Sherwood King and C. L. Carter of the University of Tennessee Experiment Station it was noted that a change in humidity had little effect on the ion density of free air. There appears to be a 12-hour cycle, but changes in temperature and humidity are not the primary cause.

Effects of Space Heating and Room Occupancy

Ion density in rooms varies according to occupancy, heat source, air contamination and circulation. Yaglou found that in unoccupied rooms the ion content of the air was not very different from that out of doors, but in occupied rooms the ion content showed a decrease depending upon the size of the room and the number of occupants.[7] Experiments conducted by the Wesix Electric Heater Company have established the change in ion density that results from use of a heater employing high temperature metal, 1,000 degrees Fahrenheit (F) and higher. Experiments by Yaglou, confirmed by others, have established the changes that result from passing air through long ducts, and measurements show that the air is not freed from ions even after it passes through filters and is washed with water.[8]

The magnitude of the charge resulting from heating and transporting air may be considerable. In one experiment the positive ion density in a large room was raised from 800 ions per cc to over 17,000, and the negative ions were reduced from 200 ions per cc to almost zero by an industrial fan-type electric heater (Fig. 2). Most of the positive ions were produced by the elements of the electric heater operating at a temperature of about 1,000 F. A few can be attributed to the presence of 19 students in the classroom. Initially, the negative ions disappeared completely, but after about 2 hours' operation there was again a small concentration. This is believed to be related to the lack of physical activity of the students and the consequent light breathing and small absorption of negative ions after 2 or 3 hours.

Similar changes, though less prominent have been observed in rooms supplied by air from ventilation systems. Change in the ion content of air supplied through ducts results from the selective absorption of negative ions because these ions have a higher mobility on the average than positive ions and are therefore lost to the walls of the duct by diffusion at a higher rate.

Aiding this natural tendency for negative ions to be separated out of air passing through ducts is the natural absorption of negative ions by metal surfaces. Massey states that in an encounter with a metal, a negative ion can lose its electron easily to the metal, provided the work function of the metal is greater than the detachment energy of the ion.[9] In this case there is an unoccupied level for the electron in the metal with energy equal to that of the level which it occupies in the ion, and the transfer will be a resonance phenomenon. Since most metal surfaces have work functions greater than the electron affinity of even fluorine it is clear that such surfaces will play a very important role in the destruction of negative ions. On the other hand, negative ions are relatively difficult to produce, particularly in connection with heating systems,[10] and it is seldom that the negative ion density of a room will exceed 1,500 ions per cc either naturally or by artificial means.

There may be one condition in which a net negative ion density is produced by air-conditioning in a conventional manner. This is when a water wash system is used and the fine spray causes the formation of many large negative ions by the well-known Lenard effect.[11] As a general rule, positive ions far exceed the negative ions in occupied air-conditioned rooms unless special precautions are taken to avoid it.

Control of Ionization

Control of atmospheric ions can be provided with any air-conditioning equipment. Unfortunately, it cannot be done by filtering alone. Some ion generation must be provided to introduce ions into the air.

The Ion-Controlled Electric Heater

The conventional electric heater using an exposed nickel-chromium element generates a preponderance of positive ions. These ions repel each other and disperse through the room. A simple suggestion to overcome this difficulty would be to design the heater with a maximum metal temperature well below the ion generation point of approximately 1,000 F.

This suggestion, however, would eliminate one of the principal advantages of an electric heater; air purification through incineration. Bacteria, mold spores, and organic dust particles such as wool, cotton etc., are destroyed when circulated through the conventional radiant-convection type of electric heater in which the air speed is approximately 180 feet per minute. The temperature of the element wire in these heaters varies from 1,100 to 1,500 F. The temperature of the ceramic supporting chimney for the coil will run 900 to 1,200 F. This is sufficient to incinerate all bacteria and organic material that go through the chimney and most of that which passes within 1 inch of the coil itself.

At the Naval Biological Laboratories, located at the Naval Supply Center in Oakland, Calif., incineration is used as the only sure means of destroying dangerous bacteria being handled in the Laboratories for experimental purposes. Various filters and aerosols were tried but the simple use of incineration was found to be by far the most effective.

The thermal force that causes precipitation on colder surfaces tends to drive particles away from a red-hot heater element unless special precautions are taken. However, this force is equally disposed on particles passing through the cylindrical chimney, and is therefore ineffective in preventing incineration of particles that enter the chimney. Near the outside of the chimney considerable improvement can be achieved by providing an electrostatic field to counter the thermal gradient directing particles away from the element.

Such an electrostatic field can also be used to control the ions coming from the heating element. In both cases the field from a negatively charged screen can be used for it is desired to collect positive ions and attract positively charged particles. The structure in its simplest form consists of a grid of 1/8-inch mesh, placed over the heating element and spaced approximately 1/8 inch from the element coil (Fig. 3). A negative potential is impressed on this grid to attract positive small ions projected from the heated coil. This negative charge will also attract small charged particles. The proximity of the grid to the coil provides sufficient heat on the grid to cause incineration at that point.

It has been discovered that an unfiltered rectified voltage applied to the grid will project the few negative ions availa-

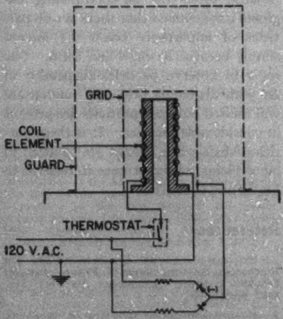


Fig. 3 (left). Ion-controlled electric heater with charged grid close to element wire

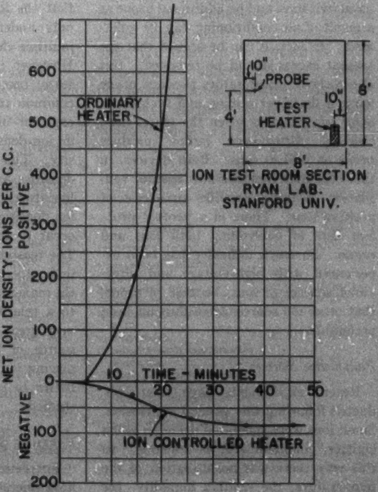


Fig. 4 (right). Comparison of room ion density characteristic of an ordinary electric heater versus an ion-controlled electric heater

ble into the airstream where they will be carried into the room. The phenomenon is not unlike that which takes place in the standard radio tube with filament, grid, and plate. At the moment that there is no potential on the grid, negative ions flow through the openings in the grid. When the negative charge again comes on the grid it forces the negative ions further away from the grid and into the airstream. Thus, a negative space charge can be created in a room. This is useful in overcoming the tendency of a positive space charge to form when the room is occupied, particularly by cigarette smokers. Fig. 4 illustrates the results achieved using this method of ion control on a radiant-convection electric heater. The addition of the grid adds very little to the cost of the heater and effectively improves air purifiction and ion control.

Disadvantages of High-Voltage and Ultraviolet Ion Sources

To provide desired ion concentrations, various ion sources are possible. The use of a thermionic source, as the hot wire of an electric heater, has been mentioned. High-voltage discharges and X-ray and ultraviolet radiation are also ion sources. The latter are limited because of the danger of producing ozone and nitrogen oxides at the energy levels required to produce ions.

In areas of air pollution we already have abnormal amounts of ozone (as found in the Los Angeles area by the Stanford Research Institute). This excess of ozone is responsible for much throat and nasal irritation and no additional ozone as a result of air-conditioning can be tolerated. In fact, it can be argued that the present excess should be removed. But electrostatic precipitators produce ozone when a flashover occurs, and such flashovers are frequent when the plates need cleaning. Ultraviolet lamps produce ozone as well as ions. Both devices, in the form that they are now used in ventilation systems, provide an excess of positive ions. There is a twofold problem here of controlling both ions and ozone. So much difficulty has been experienced with high-voltage and ultraviolet sources of ions, because of ozone, that other ion sources have thus far been preferable.

Radioactive Sources

Recent experiments have been conducted following this general conclusion. Small ions were generated using an alpha emitter, polonium *210*, as a source. In this way there was no alteration of the temperature, the relative humidity, the air velocity, or any possibility of producing ozone, oxides of nitrogen, or hydrogen peroxide. The polonium foil material chosen for this work is a carefully prepared commercial product entirely free from migration and possible radioactive contamination.[12]

The ions formed by polonium are initially small ions with high mobilities. Over a period of time some may attach themselves to particulate matter present and become intermediate-sized ions and possibly a few large ions. Before this takes place many of them will be lost by absorption on near-by objects. Thus, a constant supply of ions must be produced in order to maintain a specified density.

Convenient ion generators have been developed so that there is no longer any problem in ionizing an animal cage for experimental purposes. The same devices can be used in the home or clinically. For large forced air systems a different unit, now being developed, is preferable.

Experimental Studies

Experimental work is continuing at Stanford University and at the University of Arizona. This work is at present directed at accurately measuring ion density so as to provide automatic means of control.

ION COUNTERS

Ion counters measure not the number of ions but the number of unit charges of electricity, so that their use for the counting of ions depends on the assumption that the ions are singly charged. It is only under unusual circumstances that multiple-charged ions exist in free air.[13] However, in a device such as an electrostatic precipitator, multiple charges are common and sometimes they confuse the engineer into thinking he has obtained an ion density several times its true density. This probably explains very high densities as occasionally reported in the literature. Present indications are that ion densities above 25,000 per cc will not occur in the air of a room or building. The reasons for this are not as yet clearly understood, but the forces acting to limit the maximum density that can be achieved to a relatively low value appear to be very great. The use of radium D as a source of unipolar ions, together with placing ion generators in such a way as better to spread unipolar ions, is being studied.

PARTICLE SIZE

For determination of particle size, the light microscope is limited in its application. It possesses the advantage that the true size and shape of the particle can be observed, provided the radius of the primary particle is not less than 0.3 micron.[14] Particles between 0.1 and 0.3 micron can be detected by polarization.[15] Below this size the microscope cannot be used at all because smaller particles scatter light to an extent which depends on the inverse fourth power of the wave length of the light and on the sixth power of the radius of the scattering particle.

The electron microscope is not of much help because particles are not observed in the dispersed state and may be altered during or after sampling. This is especially true of fog droplets which are evaporated and distorted by the energy in the electrom beam.[16] The electron beam is an ionizing source itself and cannot be used in studying the behavior of charged particles.

Some work has been done by tagging very small particles with radioactive material, or by various means making them radioactive. Much of this work is classified and we can only guess at the results. In the San Francisco Bay area two major efforts are being made using this method, one at the Navy's Bacteriology Laboratory under Dr. A. P. Krueger and the other at the USNRD Laboratories at Hunters Point under the direction of Dr. E. R. Tompkins and Lt. Cdr. Royce Skow. At the New Mexico School of Mines some work has been done on "naturally" radioactive particles in air. It was learned that the observation of a major size group at 0.018 microns agrees well with previous knowledge of the dominance of intermediate atmospheric ions having mobility of about 0.02 centimeter per second per volt per centimeter[17] An indication of the presence of particles down to about 0.002 micron was detected using this method.

The importance of the difficulty of observation has not been stressed enough because often the air-conditioning engineer has assumed that there are no particles of importance below 0.1 micron simply because he can't find them. Our tools to observe particles suspended in air with diameters below 0.1 micron are still limited but such particles are present in considerable number. It is these particles which when ionized are believed to be of biological importance in relation to health and comfort.

References

1. Ionization of the Atmosphere and its Biological Effects, Lewis R. Koller. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 214, 1932, pp. 543–68.

2. Health Effects of Ions and of Ozone on

Comfort Air-Conditioning, F. W. Hutchinson. *Heating and Ventilating*, New York, N. Y., vol. 4, March 1944, pp. 76–78.

3. Atmospheric Electricity (book), J. A. Chalmers. Clarendon Press, Oxford, England, 1942, p. 29.

4. The Size-Frequency of Particles in Mineral Dust, H. L. Green. *Transactions*, The Faraday Society, London, England, vol. XXXII, 1936, p. 1091.

5. Influence of Particle Size Upon the Retention of Particulate Matter in the Human Lung, J. H. Brown, et al. *American Journal of Public Health*, New York, N. Y., vol. 40, no. 4, April 1950, pp. 450–58.

6. Control of Air Ions in Rooms, H. H. Skilling, J. C. Beckett. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 256, no. 5, 1953, pp. 423–34.

7. The Influence of Respiration and Transpiration on Ionic Content of Air of Occupied Rooms, C. P. Yaglau, et al. *Journal of Industrial Hygiene*, Baltimore, Md., vol. 15, 1933, p. 8.

8. Ionization of Air in an Air-Conditioned Building. *Nature*, London, England, vol. 142, Nov. 26, 1938, p. 956.

9. Negative Ions (book), H. S. W. Massey. University Press, Cambridge, England, 1950, p. 89.

10. Climate Control Through Ionization, T. L. Martin, Jr. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 254, 1952, p. 267.

11. Terrestrial Magnetism and Electricity (book), J. A. Fleming. Dover Publications, New York, N. Y., 1949, p. 169.

12. Production of Unipolar Air with Radium Isotopes, Thomas L. Martin, Jr. *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 771–76.

13. Fundamental Processes of Electrical Discharge in Gas (book), L. B. Loeb. John Wiley and Sons, Inc., New York, N. Y., 1947, p. 49.

14. Dispersed Systems in Gases: Dust, Smoke and Fog, G. M. B. Dobson. *Transactions*, The Faraday Society, London, England, vol. XXXII, 1936, p. 1150.

15. The Preparation, Collection, and Measurement of Aerosols, Victor K. LaMer. *Proceedings*, First National Air Pollution Symposium, Stanford Research Institute, Los Angeles, Cal., 1949, p. 6.

16. Handbook of Aerosols. U. S. Atomic Energy Commission, Washington, D. C., 1950, "Particle Size and Size Distribution," David Sinclair, p. 98.

17. Natural Radioactivity as a Tracer in the Sorting of Aerosols According to Mobility, M. H. Wilkening. *Review of Scientific Instruments* New York, N. Y., vol. 23, no. 1, p. 15.

No Discussion

Characteristics of Aircraft A-C Generators

L. J. STRATTON
ASSOCIATE MEMBER AIEE

L. W. MATSCH
FELLOW AIEE

THE performance of 208/120-volt 400-cycle constant-frequency aircraft electric power systems under abnormal conditions resulting from synchronizing disturbances, loss of field excitation, and short circuits is described. As a basis for this study, tests were made on 40-kva alternators, to determine their electrical parameters. The values of these parameters expressed in per unit are listed in Table I.

Alternator Parameters

The alternator in its simpler aspects may be considered an ideal voltage source feeding its load through a variable impedance. Under balanced conditions only one phase of the alternator need be considered. Figs. 1 and 2 show the equivalent circuit and phasor diagram of a nonsalient pole alternator, and Fig. 3 shows the phasor diagram of a salient pole alternator. From Figs. 1 and 2 it can be shown that the equation for the terminal voltage of a nonsalient pole alternator is

$$V = E - I(R + jX_d) \tag{1}$$

and from Fig. 3 that of a salient pole alternator is

$$V = E - RI - jX_q I_q - jX_d I_d \tag{2}$$

where

V = terminal voltage
E = internal voltage
R = armature resistance
I = armature current
X_d = direct-axis synchronous reactance
X_q = quadrature-axis synchronous reactance
I_d = direct-axis armature current
I_q = quadrature-axis armature current

The steady-state parameters of alternators are determined from the following characteristics: no-load saturation curve, and zero power-factor saturation curve (two points). These characteristics are shown graphically in Fig. 4. From the right triangle shown in this figure, the Potier reactance and demagnetizing factor were found to be 0.11 per unit and 1.43 per unit respectively. The synchronous impedance determined from these two characteristics ranges from its unsaturated value of 1.5 per unit to a saturated value of 0.15 per unit at ceiling excitation (5.75 per unit). Fig. 5 shows the zero power-factor volt-ampere characteristic of the 40-kva alternator. In salient-pole alternators, the air gap is not uniform, and quadrature-axis and direct-axis synchronous reactances are used. The unsaturated quadrature-axis synchronous reactance was found to be 0.91 per unit from a slip test at approximately half-rated voltage.

For analyzing faults or other unbalanced operation of alternators, the positive, negative, and zero sequence quantities are used. The manner in which these are used is shown in textbooks on symmetrical components.[1-3] The positive-sequence impedance was determined from oscillographic records of the armature currents immediately after the occurrence of a 3-phase short circuit,[3] and was found to be $(0.024 + j0.15)$ per unit. In conventional power systems using high-speed circuit breakers, the subtransient impedance is taken as the positive-sequence impedance. In systems where slower speed circuit breakers are used, the transient impedance is used. Although the 40-kva alternators tested had damping grids, the subtransient and transient reactance were substantially equal, as shown in Table I. The nega-

Table I. Alternator Parameters

Quantity	Symbol	Per Unit Value
Unsaturated synchronous reactance	X_{du}	1.50
Short-circuit ratio	SCR	0.90
Potier reactance	X_p	0.11
Armature demagnetizing factor		1.43
D-c resistance	R	0.024
Quadrature-axis synchronous reactance	X_q	0.91
Direct-axis transient reactance	X_d'	0.15
Direct-axis subtransient reactance	X_d''	0.15
Quadrature-axis subtransient reactance	X_q''	0.54
Negative-sequence impedance	Z_2	0.14+ j0.22
Zero-sequence impedance	Z_0	0.024+ j0.042

Paper 52-319, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Middle Eastern District Meeting, Toledo, Ohio, October 28–30, 1952, and represented at the 1954 Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted August 6, 1952; made available for printing September 5, 1952.

L. J. Stratton and L. W. Matsch are with the Illinois Institute of Technology, Chicago, Ill.

The material presented in this paper is based on a study sponsored by the Electrical Branch, Equipment Laboratory, Wright Air Development Center.

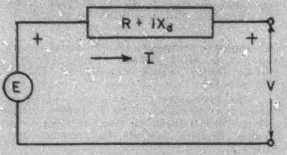


Fig. 1. Equivalent circuit of a nonsalient pole alternator

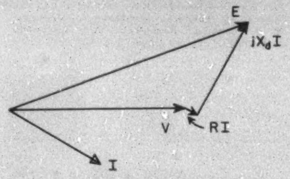


Fig. 2. Phasor diagram of a nonsalient pole alternator

tive-sequence impedance was determined to be $(0.14+j0.22)$ per unit with reduced field excitation and with the alternator driven at synchronous speed using connections shown in Fig. 6. With full excitation, the negative-sequence reactance is probably lower because of magnetic saturation. Since the neutrals of aircraft alternators are grounded, a path for zero-sequence current is provided. The zero-sequence impedance Z_0 was found to be $(0.024+j0.043)$ per unit by applying single-phase voltage to all three phases in series. The values of impedance, reactance, short-circuit ratio, and demagnetization factor are listed in Table I.

Synchronizing Disturbances

Improper synchronizing of alternators may result in dips in the system voltage and a large interchange of power between the incoming alternator and the system. If the effect of saliency is neglected and values are substituted for δ, the torque angle, into the power expressions derived in the Appendix, along with typical values of resistance and reactance, the loci of the power phasor are two families of ellipses as shown in Fig 7. The dotted ellipses represent the loci of the power phasor for each of the $(n-1)$ alternators already on the bus and the solid ellipses are the loci for the incoming alternator. For a 2-alternator system, $n=2$, the two ellipses become one straight line through the origin. For an alternator synchronized to an infinite bus, $n=\infty$, the ellipse for the incoming alternator is a circle, and that for any one of the other alternators a point at the origin. Fig. 7 shows that in a system of three or more alternators, the maximum reactive power which an incoming alternator may deliver is small compared to that which it may draw from the system. However, large swings of real power may result depending upon the value of δ at the time of synchronizing.

Fig. 8 shows oscillograms of field current, armature current, voltage, and power when two alternators, designated as *A* and *D*, were synchronized out of phase. An overrunning clutch is incorporated in the coupling between the alternator and its constant speed drive, thereby permitting ease in sychronizing. If $n=2$ is substituted in equations 8 and 9, the ratio of P to Q is $X_d'/R=3.5$, where X_d' is the transient reactance and R the resistance of the armature and leads to the bus ($X_d'=0.15$ and $R=0.043$). From the oscillograms of Fig. 8, the ratio P to Q is found to be 3.7. Equation 7 for the bus voltage shows that the power oscillations are accompanied by a reduction in the bus voltage. For the most severe condition of synchronizing 180 degrees out of phase, the bus voltage goes to zero, and the armature current from equation 6 is 6.5 per unit for the 2-alternator system. For an infinite system the bus voltage is unaffected but the current in the incoming alternator is 13.0 per unit or twice the available 3-phase short-circuit current. Such an extreme transient current as 13.0 per unit would probably cause the exciter to reverse its polarity which causes the stabilizing transformer in the voltage regulating circuit to malfunction.[4]

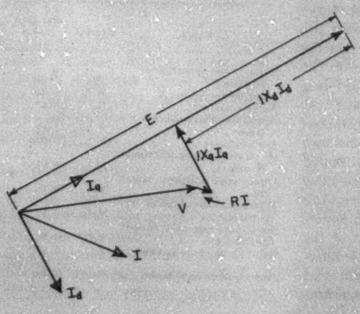


Fig. 3 (left). Phasor diagram of a salient pole alternator

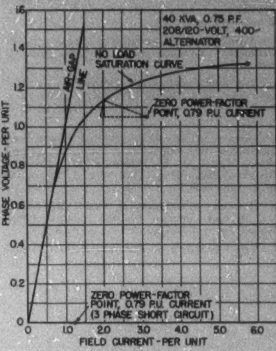


Fig. 4. Alternator characteristics

Loss of Field Excitation

An alternator may lose its field excitation as a result of an open circuit or complete short circuit in either its own field circuit or in the exciter field circuit. A salient-pole alternator without field excitation can carry some real load without losing synchronism when operating in parallel with other synchronous machines. The expressions for the real power delivered to, and the reactive power drawn from the system for an alternator having zero field excitation are

$$P=\frac{X_d-X_q}{2X_dX_q}|V|^2\sin 2\delta \quad (3)$$

$$Q=\frac{|V|^2}{2X_dX_q}[X_d+X_q-(X_d-X_q)\cos 2\delta] \quad (4)$$

where $|V|$ is the terminal voltage, and δ is the torque angle. If the prime mover input to the unexcited alternator exceeds the maximum electrical power indicated by equation 3 (plus the rotational losses), the alternator will be driven above synchronous speed. In this sense synchro-

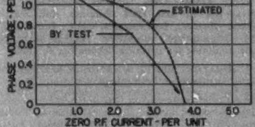


Fig. 5 (right). Volt-ampere characteristics for ceiling excitation and zero power-factor current

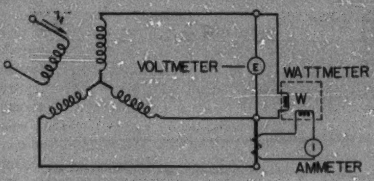


Fig. 6 (left). Connections for determining negative - sequence impedance

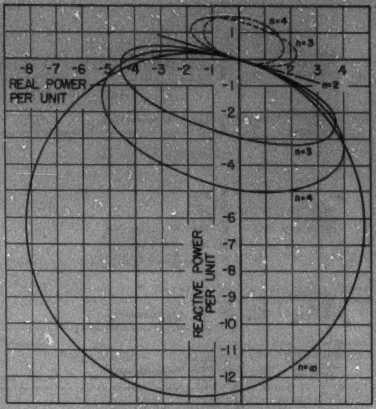


Fig. 7 (right). Power relationship on synchronizing

nous speed is considered proportional to the system frequency whether the frequency is normal or abnormal. Neglecting induction generator effects, the equation for the real power under this condition would be expressed by

$$P=\frac{X_d-X_q}{2X_dX_q}|V|^2\sin 2[2\pi(f_1-f)t+\delta_0] \qquad (5)$$

where f is the system frequency (corresponding to synchronous speed), f_1 is the frequency corresponding to the speed of the unexcited alternator, and δ_0 is the value of δ at the instant when the field excitation became zero.

The effect of opening the field circuit of one alternator when operating in parallel with another is shown in the oscillograms of Fig. 9. The conditions under which the oscillograms in Fig. 9 were obtained just exceeds the limiting condition for which synchronism is maintained when the field of an alternator is opened since the slip speed is low (2 per cent). The alternator with the open field is driven about two revolutions per second faster than the other alternator Fig. 9 shows that the system voltage oscillates between 0.9 and 0.6 per unit after the field is opened. The reduction is due to reduced excitation of the healthy alternator caused by the reactive load equalizer circuit.[5] As the load used in the tests is passive, its magnitude varies approximately as the voltage squared. To approximate the load for the condition of Fig. 9 the average of 0.6 and 0.9 is taken for the voltage. Hence, the real power is

$$P=0.30\times(0.75)^2=0.17 \text{ per unit}$$

This limiting value of real power can be checked by the use of equation 3 using the values $X_d=1.50$ per unit, $X_q=0.91$ per unit, and $|V|=0.90$ per unit.

$$P_{\max}=\frac{1.50-0.91}{2\times1.50\times0.91}(0.9)^2=0.175 \text{ per unit}$$

The real power limit for sustained synchronism with an open field is approxi-

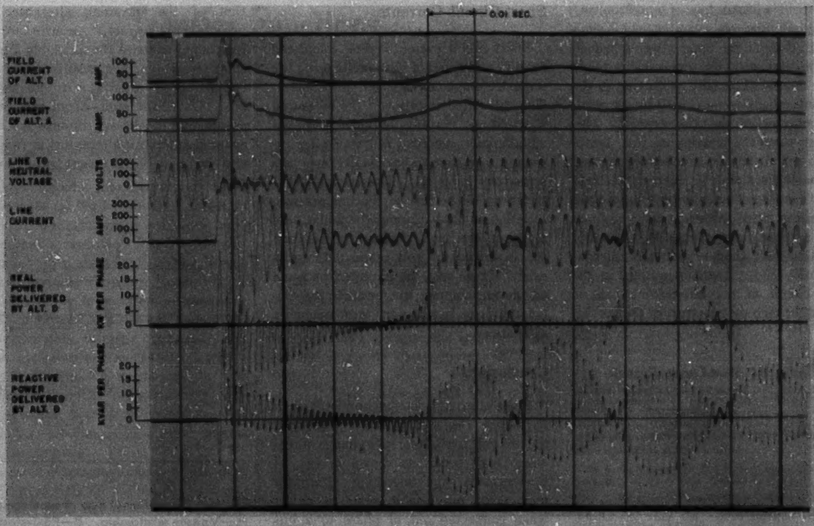


Fig. 8. Effect of synchronizing two alternators out of phase

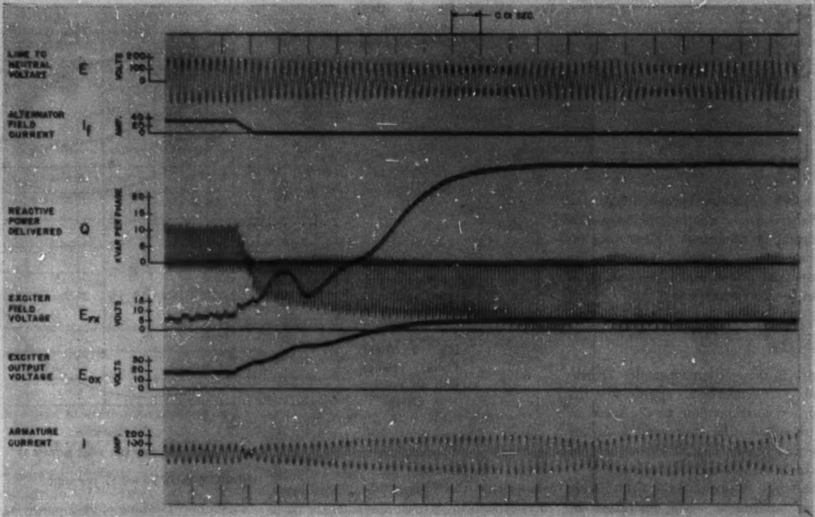


Fig. 9. Effect of opening field of one of two alternators while operating in parallel. System load 36.0 kva, 0.66 power factor

mately 0.17 per unit for alternators with the parameters listed. However, for a system with more than two alternators, the reduction in the system voltage is less, and as a result a larger real power load can be delivered by the alternator with the open field without loss of synchronism.[5] The effect of an open exciter field is similar to that of an open alternator field except that with the open exciter field there is a small amount of current in the alternator field circuit caused by the residual magnetism in the exciter. This means that the maximum real power which an alternator can deliver at synchronous speed with an open exciter field is somewhat greater than that for an open alternator field. When synchronism is lost, with an open exciter field the slip speed is lower than for an open alternator field, since current is induced in the field circuit of the alternator which is closed through the armature of the exciter.

Table II. Short-Circuit Currents

Type of Fault	Measured Current, Per Unit	Calculated Current, Per Unit
Single-line-to-ground	5.7	6.1
Line-to-line	5.0	4.1
3-phase	6.5	

The line impedances are, per unit: $Z_1=0.019+j\,0.005$ per unit; $Z_2=0.019+j\,0.005$ per unit; $Z_0=0.021+j\,0.014$ per unit.

When an alternator loses its field excitation when operating in parallel with other alternators, it draws large amounts of reactive power from the system for its excitation.

Short-Circuit Currents

Short-circuit currents were calculated for different types of solid faults using the sequence impedances listed in Table I. The calculated results and the measured short-circuit currents obtained from oscillographic records of single-line-to-ground and line-to-line faults are presented in Table II. The sequence impedances of the line to the bus are also included in Table II since the faults were made at the bus. Differences between the experimental and calculated results exist. These may be expected as some of the parameters, upon which the calculations are based, vary with magnetic saturation.

Conclusions

If an alternator is synchronized when its voltage is appreciably out of phase with the bus voltage, a large interchange of real power may occur, as shown in Fig. 7. Furthermore, the incoming alternator may draw appreciable reactive power when connected out of phase to a bus supplied by two or more alternators, whereas the maximum reactive power it can deliver to the system is relatively small. A severe reduction in the bus voltage may result, depending on the number of alternators on the system when the angle between the bus voltage and that of the incoming alternator approaches 180 degrees. An armature current equal to twice the available 3-phase short-circuit current results from synchronizing an alternator to an infinite bus. Under the stress of an emergency, an operator might synchronize machines as small as aircraft alternators at random, and the phase displacement between volt-

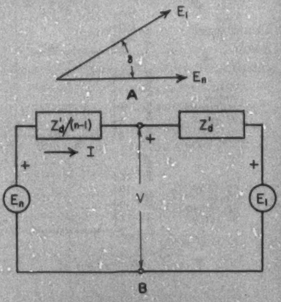


Fig. 10. Conditions at time of synchronizing. A—Voltage relationship. B—Equivalent circuit

ages may have any value from zero to 180 degrees. Automatic synchronizing is desirable, and methods should be developed which would keep the phase displacement small enough so that voltage dips and power swings are negligible.

When an alternator loses its field excitation while operating in parallel with other alternators, it draws a large amount of reactive power from the system and causes the system voltage to be subnormal and even to oscillate for some loads. The alternator with the open field circuit should be promptly disconnected from the system.

Appendix I. Power Equations

Consider an unloaded bus supplied by $(n-1)$ alternators and to which an alternator (alternator 1) is connected when its voltage leads that of the bus by the angle δ. Fig. 10 shows a diagram of the equivalent circuit of the bus and the incoming alternator, and also a phasor diagram which shows the voltage $\boldsymbol{E}_1$ of the incoming alternator to lead the bus voltage $\boldsymbol{E}_n$ by the angle δ. Let the voltages be expressed by

$$\boldsymbol{E}_n = |E| + j0$$

$$\boldsymbol{E}_1 = |E|(\cos\delta + j\sin\delta)$$

If the effect of saliency is neglected, the current is expressed by

$$I = \frac{|E|(1-\cos\delta - j\sin\delta)}{Z_d' + Z_d'/(n-1)} \qquad (6)$$

where

n = total number of alternators including incoming alternator
$Z_d' = R + jX_d'$
R = resistance of the armature plus that of the leads to the bus
X_d' = transient reactance of the alternator plus the reactance of the leads to the bus

The terminal voltage $\boldsymbol{V}$ in terms of $|E|$ and δ is

$$\boldsymbol{V} = E_n - \frac{Z_d'}{n-1} I = |E|\left[1 - \frac{1-\cos\delta}{n} + j\frac{\sin\delta}{n}\right] \qquad (7)$$

The equations for the real and reactive power delivered by the incoming alternator are respectively

$$P_1 = -(n-1)\frac{|E|^2}{n|Z_d'|^2}\left[R\left(\frac{n-2}{n}\right)\times (1-\cos\delta) - X_d'\sin\delta\right] \qquad (8)$$

$$Q_1 = -(n-1)\frac{|E|^2}{n|Z_d'|^2}\left[X_d'\left(\frac{n-2}{n}\right)\times (1-\cos\delta) + R\sin\delta\right] \qquad (9)$$

Similar power equations may be obtained for the alternators originally on the bus prior to synchronizing.

References

1. SYMMETRICAL COMPONENTS (book), C. F. Wagner, R. D. Evans. McGraw-Hill Book Company, Inc., New York, N. Y., 1933.

2. THE TRANSMISSION OF ELECTRIC POWER (book), W. A. Lewis. Illinois Institute of Technology, Chicago, Ill., lithoprinted edition, vol. II, 1948.

3. ELECTRICAL TRANSMISSION AND DISTRIBUTION REFERENCE BOOK. Westinghouse Electric Corporation, East Pittsburgh, Pa., 4th ed., 1950.

4. EXCITER POLARITY REVERSALS IN VOLTAGE-REGULATED AIRCRAFT ALTERNATORS, R. P. Judkins, H. M. McConnell. *AIEE Transactions*, vol. 71, pt. III, Jan. 1952, pp. 275–83.

5. POWER EQUALIZER SYSTEMS FOR AIRCRAFT ALTERNATORS, John A. Granath, A. K. Hawkes. *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 209–17.

◆

No Discussion

Self-Oscillation Method for Measuring Transfer Functions

JOHN C. CLEGG
ASSOCIATE MEMBER AIEE

L. DALE HARRIS
MEMBER AIEE

THE person analyzing or synthesizing a closed-loop control system often has need of sinusoidal transfer function data. If the data are to be obtained directly, ordinarily a measurement is made of both phase and magnitude of the sinusoidal response to a sinusoidal driving force. A low variable-frequency generator of sufficient power is required, as well as a means of measuring phase and magnitude of the sinusoidal response quantity. The phase-angle measurement is often troublesome and rather inaccurate at the low frequencies which are commonly employed in the closed-loop control system. Sometimes nonlinearities further aggravate the problem.

Sinusoidal transfer function data may be observed without the use of a separate sinusoidal generator or driver if the system is allowed to oscillate and thus supply its own driving force; furthermore, the necessity of the measurement of magnitude and phase of a sinusoid is eliminated. The frequency of oscillation must be measured, using the self-oscillation method, but this quantity is easily and accurately observed. Other advantages of the self-oscillation method will become apparent as the idea is developed.

The advisability of intentionally allowing a closed-loop system to oscillate may be questioned, because the amplitude of oscillation may be destructive to the system.

It is demonstrated in this paper that the amplitude of oscillation can be held easily to a safe value and, furthermore, varied over a wide range to yield transfer function data as a function of amplitude as well as frequency. Thus the method is applicable to the describing function[1] as well as to the linear point of view.

Self-Oscillation Method

To find G, the sinusoidal transfer function of the unknown system, this system is cascaded with an auxiliary system called the calibrated system. See Fig. 1. The transfer characteristics of the calibrated system are adjustable and known.

The closed-loop transfer function of the composite system is given by equation 1

$$\frac{\theta_o}{\theta_i} = \frac{G_T}{1+G_T} \qquad (1)$$

where G_T is the transfer function of the composite system. When the system is in steady-state oscillation, θ_o is finite and θ_i is equal to zero; the denominator of equation 1 must be equal to zero

$$1 + G_T = 0 \qquad (2)$$

but

Paper 54-284, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing April 23, 1954.

JOHN C. CLEGG and L. DALE HARRIS are with the University of Utah, Salt Lake City, Utah.

This paper is an outgrowth of Research Contract No. *DA-36-039-SC-42621*, sponsored by the U. S. Signal Corps. Grateful acknowledgment is extended to the personnel of the servomechanism section of Evans Signal Laboratories, Belmar, N. J.

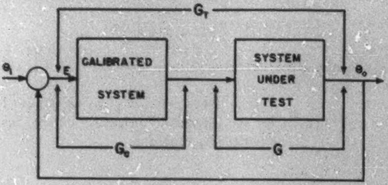


Fig. 1 (left). Block diagram showing self-oscillation method

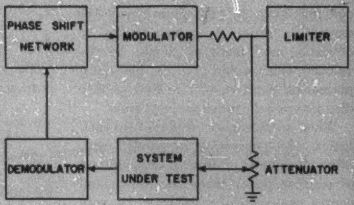


Fig. 2 (right). Some details of the calibrated system

$$G_T = GG_c \qquad (3)$$

therefore

$$G = -\frac{1}{G_c} \qquad (4)$$

If the composite system is allowed to oscillate, the unknown transfer function G is related to the known or calibrated transfer function G_c as shown by equation 4. The calibrated system is given phase-shift and attenuation characteristics which are adjustable and are readily determined for each and every frequency of oscillation. To control the amplitude of oscillation, the calibrated system also contains a saturating element whose transfer function is readily determined through the application of describing-function concepts.[1] The frequency of oscillation is changed primarily by a change of the phase-shift characteristics of G_c. The amplitude of oscillation is controlled primarily by the level of saturation in the saturating element. However, if significant nonlinearities are contained in the system under test, then the level of saturation will have some effect on the frequency of oscillation and the phase control will have some effect on the amplitude of oscillation. The adjustable attenuation characteristic of the calibrated system in general must include an amplifier to achieve both positive and negative attenuations. The attenuation level is adjusted for a stable oscillation of a magnitude such that the signal does not swing excessively into the saturation region of the saturating element. In this manner a minimum of harmonic distortion is introduced by the calibrated system. By adjusting attenuation, saturation level, and phase of the calibrated system, transfer function data for the system under test are obtained.

Fig. 2 shows an arrangement used for testing the validity of this theory. In this case the signal in the servo system under test was in the form of a suppressed carrier wave typically found in an a-c servo system. It was considered best to demodulate the wave in order to use the simpler phase-shifting networks having transfer functions which are readily found. The limiter was a typical vacuum-tube circuit having the characteristics displayed in Fig. 3.

Fig. 4 shows a locus of a particular transfer function obtained by the self-oscillation method. This locus was observed at one particular signal amplitude at the output, but the composite system could be made to oscillate at each frequency shown on this locus through a range of amplitudes such that the ratio of maximum to minimum amplitude was approximately 8. Ordinary laboratory components were used in assembling the so-called calibrated system for this test. It is believed that with moderate time and expense a calibrated system could be designed so that for each frequency the transfer function of the calibrated system or perhaps the system under test could be read directly from a set of dials.

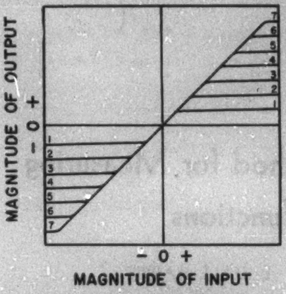


Fig. 3. Typical characteristics of a vacuum-tube limiter. Limit level is set at a low value for curve 1 and at a high value for curve 7

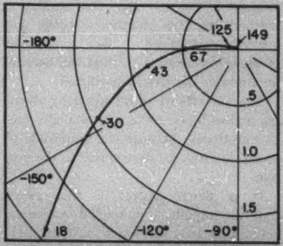


Fig. 4. Locus of a transfer function of a typical instrument servo system, measured by self-oscillation method for a constant amplitude. Locus plotted in a polar co-ordinate system. Numerals adjacent to locus are frequencies in radians per second

Conclusions

The self-oscillating method can be used successfully to obtain the locus of the sinusoidal transfer function of a closed-loop system. For nonlinear systems the method conveniently yields data as a function of amplitude as well as of frequency. The harmonic distortion arising from the nonlinearities does not interfere with the measurements. When this method is used, the need for a separate driving force, sinusoidal or otherwise, is eliminated. Transient, sinusoidal phase, and sinusoidal amplitude measurements are not made. Where more conventional methods are applied to some of the more complicated systems, the engineer is forced to close the loop on an unstable system, not knowing the locus of the transfer function. Often, this system oscillates with amplitudes of destructive proportions. The engineer is still unable to arrive at a stabilizing method except by cut-and-try processes. The self-oscillation method is well adapted to this situation. By inserting the previously described calibrated block, the amplitude of oscillation is held within safe limits while pertinent transfer function data are observed on the otherwise violently unstable system. Now the engineer may base the design of a compensation system on the transfer function data thus obtained.

Reference

1. SINUSOIDAL ANALYSIS OF FEEDBACK-CONTROL SYSTEMS CONTAINING NONLINEAR ELEMENTS, E. Calvin Johnson. *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 169–81.

◆

No Discussion

Rebuilding San Francisco's Transit System

F. L. ReQua
MEMBER AIEE

SAN FRANCISCO is situated on the northern extremity of a peninsula, surrounded on three sides by water. Its population of 800,000 is confined within a rectangular area of 43 square miles. It is one of the hilliest cities in the United States, with land elevations varying between sea level and 930 feet. North of Market Street and the diagonal line bisecting the city, the streets are laid out in a substantially north-south, east-west grid pattern; see Fig. 1. South of this line the street pattern is made up either of smaller grids, oriented in many different directions, or of contour streets in and around the hillier sections.

The first street railroad in San Francisco, founded in 1857, was operated by steam but was later converted to horsecar operation. The first cable line went into operation in 1873. In 1893 12 of the 17 lines then in operation were consolidated into the Market Street Railway Company; 9 years later this company and all but three of the independent companies were further consolidated into the United Railroads of San Francisco.

Prior to 1906 most of the lines were cable-operated. Damage to the cable slots and machinery as a result of the earthquake and fire in that year was so extensive, however, that advantage was taken of the opportunity thus afforded to convert extensively from cable to electric lines; from 1906 to 1910 the United Railroads spent large sums on such a program.

In 1912 the Municipal Railway of the City and County of San Francisco went into operation in compliance with the general policy laid down in the City Charter that utilities be municipally owned and operated. The first line was electric, running on tracks which by mid-1913 extended from the ferries westerly on Geary Street to the beach and park. The necessity of making preparations to handle the great number of visitors who were expected to attend the Panama Pacific International Exhibition in 1915 greatly accelerated the development of the Municipal Railway.

In 1921 the United Railroads of San Francisco was reorganized and refinanced, again under the name Market Street Railway Company, and shortly thereafter the Municipal Railway commenced efforts to acquire these properties. However, it was not until 1944 that these efforts were consummated, and the two systems, of substantially equal size, were consolidated. Still later, in 1952, the Municipal Railway acquired the last independent line in the city, the California Street Cable Railroad Company.

Equipment and trackage acquired through purchase of the Market Street Railway was very old and had become hopelessly dilapidated. The municipal system at the time of consolidation was by contrast well kept up, although much of the equipment was very old. About \$3,000,000 was spent on rehabilitation immediately following consolidation, but it was evident that considerably more would have to be invested to secure an adequate, modern transportation system for the city. Accordingly, a \$20,000,000 bond issue was approved by the electorate in 1947. The principal expenditures under this program, now substantially completed, are listed in Table I. Tables II and III show the character and extent of route facilities and of passenger-carrying ability before and after rehabilitation of the system.

Power Supply

In conjunction with its water supply facilities, the City of San Francisco owns and operates an 85,000-kw hydroelectric plant fed from waters stored in the Hetch Hetchy reservoir in Yosemite National Park. Power from this plant is transmitted 100 miles over city-owned transmission lines to the east side of San Francisco Bay, and from there a further 35 miles into the city over transmission lines of the Pacific Gas and Electric Company. The Municipal Railway purchases all its power from the city's Hetch Hetchy system. The resulting very favorable energy rate of about $1^{1}/_{8}$ cents per kilowatt-hour at the 600-volt d-c busses led to the adoption of electric propulsion on a very wide scale in the rehabilitation planning.

Rolling Stock

Streetcars

Of the 440 streetcars acquired from the Market Street Railway all were of wood-body construction and were of little value. Furthermore, most of the rails were in very poor condition. These considerations, together with the fact that 1-man streetcar operation is illegal in San Francisco by reason of an ordinance enacted in 1935, led to the economic necessity of substituting rubber-tired vehicles in lieu of rail vehicles wherever possible.

In consequence, only four of the electric rail lines were retained: two that operate through the 2-mile tunnel under Twin Peaks and through the adjacent 3/4-mile Sunset Tunnel and two others, the Geary Street and the Church Street lines. Unless the 2-man streetcar ordinance is repealed it is probable that the latter two lines will be converted to trolley coach in the future. If so, the only remaining rail lines in San Francisco would be those through the tunnels, which are too narrow for safe bus or trolley-coach operation, and for sentimental reasons, some or all of the present cable lines.

Of the 185 electric rail cars now in service, 145 dating back to 1912 have been retained from the original Municipal Railway fleet. These are 2-man double-end steel-body cars weighing about 50,000 pounds and with seating capacity of 50.

Fifteen cars, some dating back to 1939, are of President's Conference Committee (PCC) construction modified for 2-man double-end operation. They weigh about 40,000 pounds and seat 60. These cars have four doors, two on opposite sides at each end. The conductor's seat is located opposite the rear entrance door and the operator's seat is at the front on the left side. The first five of these cars have air compressors and manually operated controllers at each end; the remaining ten are of the all-electric type with foot-operated controllers at each end.

In 1951 25 more all-electric PCC-type cars were purchased; these are for 2-man single-end operation with center entrance door and front exit-only door. They weigh about 37,000 pounds each. The conductor's position is a seat across from the center door, but the body design of these cars is such that they can be readily converted to 1-man operation. Because the existing car houses are designed for double-end car operation, these single-end cars often have to be backed several hundred feet. Therefore, they have been equipped with a second trolley pole, provided with carbon-shoe collectors,

---

Paper 54-253, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 11, 1954; made available for printing April 26, 1954.

F. L. ReQua is with the Public Utilities Commission, City and County of San Francisco, San Francisco, Calif.

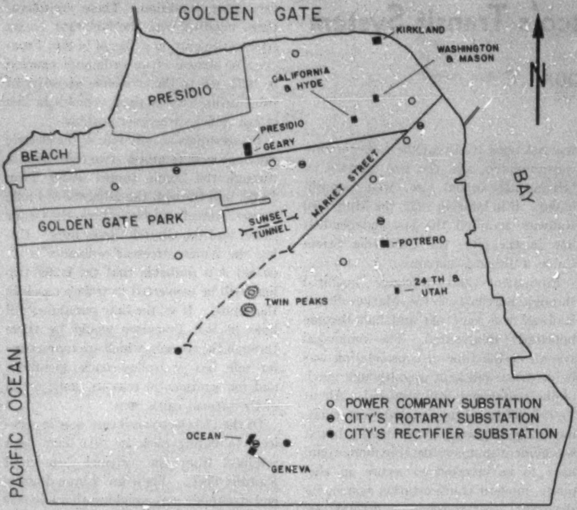


Fig. 1. Map of San Francisco

as are all other trolley poles on the Municipal Railway.

TROLLEY COACHES

Immediately following consolidation the city possessed a total of 18 trolley coaches; nine of these were 40-passenger vehicles which it had owned initially, and nine were 37-passenger 2-motor vehicles which it acquired with the Market Street property. New 44- and 48-passenger vehicles were purchased as rapidly as storage and overhead trolley facilities could be made available until today the city possesses a fleet of 398 trolley coaches.

The topography and population distribution of San Francisco is such that severe grades are encountered on many of the new trolley coach routes—up to 19 per cent on some. Because of this a great deal of thought was given to the design of an emergency braking system in addition to the conventional hand brake and dynamic brake operating on the drive shaft and air brakes operating on the front and rear wheel drums. The goal was a means for adequately decelerating a 48-passenger coach, on 20-per-cent grade with full standee load, in the event of failure of any or all of the elements of the conventional braking system.

Table I. Municipal Railway Expenditures for Rehabilitation, 1944 to 1954

Item	Amount, Dollars
Rolling stock	
Streetcars, 40 at $30,500 (average)	1,220,000
Trolley coaches, 380 at 19,100 (average)	7,258,000
Motor busses, 373 at 15,400 (average)	5,759,000
Miscellaneous work	86,000
Total	14,323,000
Storage facilities	
Trolley coaches, 2	1,248,000
Motor busses, 2	1,134,000
Total	2,382,000
Trolley system	
Overhead street construction (103.3 miles of negative conductor)	2,510,000
Overhead in two storage yards	232,000
Substations, 2 (2,000 kw and 750 kw)	288,000
Total	3,030,000
Track	
Reconstruction	1,893,000
Removal*	623,000
Total	2,516,000
Buildings, structures, and miscellaneous	636,000

* This item includes cost of removing the two outer tracks on Market Street and both tracks from Mission Street. All other track removal has been performed by another city department under a separate bond issue for street rehabilitation.

On some of the earlier vehicles such an auxiliary braking means was brought into action by forcing the hand-operated reverse lever through a normally restraining shear pin into an "emergency" position. This connected a resistor directly across the motor terminals by a mechanically actuated switch and so provided a single-valued dynamic braking position. One of the principal objections to this scheme was the fact that breakage of a rear axle would render it inoperative.

Later coaches were equipped with a supplemental braking system consisting of an auxiliary air tank mounted close to the rear wheels and connected by short pipes directly to the rear brake cylinders. This system is pneumatically isolated from the main air system by reduction orifices and check valves so that in event of loss of air in the main piping by reason of breakage or other cause, enough pressure will remain in the auxiliary system to bring the vehicle to a stop several times under the most adverse conditions likely to be encountered. It is actuated either by the dead-man pedal or failure of the battery control power.

Another feature incorporated in later coaches is a push button adjacent to the operator which, when depressed, increases the setting of the current-limiting relay and prevents the control from notching into the high-speed weak-field position. This enables the vehicles to go up the steepest grades at relatively high speed even when carrying a full standee load. To prevent overheating under these conditions the main resistors are of heavy duty type.

Much thought was also given to minimizing the possibility of electric shock to passengers boarding trolley coaches. Because the negative bus of the power system is connected to the rails the negative

Table II. Route Facilities Before and After Rehabilitation

Item	1944	1954
Miles of single track (includes cable cars)	280	56
Miles of trolley coach negative contact conductor	17	128
Number of trolley coaches	18	398
Number of motor busses	165	467
Total rubber-tired vehicles	183 (20%)	865 (77%)
Number of rail cars (includes cable cars)	716 (80%)	252 (23%)
Total vehicles	899	1,117

Table III. Weekday Peak P.M. Passenger-Carrying Potential of Municipal Railway System Before and After Rehabilitation

Type of Vehicle	1944			1954		
	Peak P.M. Vehicle Demand	Loading Capacity per Vehicle*	Total Riding Units	Peak P.M. Vehicle Demand	Loading Capacity per Vehicle*	Total Riding Units
Streetcar	538	100	53,800	125	100	12,500
Trolley coach	18	58	1,044	332	68	22,576
Motor bus	152	55	8,360	391	66	25,806
Cable car†	32	71	2,272	39	70	2,730
Total	740		65,476	887		63,612
				Calendar 1945		Calendar 1953
Revenue passengers				252,000,000		157,100,000
Approximate total riding units available for p.m. peak per million annual revenue passengers				260		405
Relative values of above				100%		156%

* Loading capacity per vehicle based on:
2.0 times seating capacity for streetcars.
1.5 times seating capacity for trolley coaches and motor busses.
65 passengers for Powell cable cars.
75 passengers for California cable cars.

† Includes 20 vehicles of the then independent California Cable Railway.

trolley coach conductors are necessarily grounded at the substations. Hence, if the insulation resistance between the 600-volt equipment and coach body becomes impaired, a person in contact with the coach body and the ground can receive an electric shock, particularly in wet weather. To minimize this possibility particular attention was given to electrical insulation design throughout the vehicle. All 12-volt conductors are insulated for 600 volts and all 600-volt conductors are insulated with neoprene-sheathed rubber without braid. In addition to these and other structural precautions, the insulation of the vehicle is automatically tested during each passage through the carhouse.

The latest trolley coaches and streetcars are all equipped with a dead-man pedal arranged to shut off power and apply the emergency brakes when the pedal is in either the fully released or fully depressed position. The weight of the operator's foot and leg depresses the pedal to the normal operating position at rest against a resilient stop. It can be moved into the fully depressed position only by exercise of considerable muscular effort, thus taking advantage of the instinctive muscular reflex tendency to brace the body against an impending collision.

The matter of road clearance was also given considerable study. Unusual grades encountered on many San Francisco streets require correspondingly abrupt transition in the cross streets. It was therefore necessary to specify adequate clearance over vertical curves of a 200-foot radius, crown or depression, with fully loaded vehicle and with allowance for passage over possible small objects or irregularities on the road.

Year-round temperatures in San Francisco are such that car and coach body heating is not required. Therefore no provision for artificial heating was made other than a small electric heater located at the operator's feet.

MOTOR BUSSES

At the present time the city owns 467 motor busses, all powered by liquid fuel. Of these, 33 owned by the Municipal Railway before the consolidation and 45 acquired from the Market Street Railway are over 12 years old. After consolidation 16 used Diesels were acquired from the War Assets Corporation at a time when new equipment was not obtainable because of the war. The remaining 373 gasoline-powered busses, now 6 to 9 years old, were purchased new as a part of the rehabilitation program. They are for the most part 44-passenger vehicles.

Storage Facilities

STREETCAR STORAGE

At the time of consolidation the city acquired six car houses from the Market Street Railway, all of which were dilapidated and some of which were considered unsafe for further use. All but two have been disposed of, the Geneva car house and the cable car house at Washington and Mason Streets. The Geneva car house, with its ladder track system, has been reconditioned and serves, together with the Municipal Railway's original reinforced concrete Geary Street car house, to store the 185 electric streetcars now operated.

The Washington-Mason cable car house and the cable car house acquired from the California Street Cable Railway Company, located at California and Hyde Streets, together house the 67 cable cars now owned by the Municipal Railway.

TROLLEY COACH STORAGE

Two storage yards, Potrero and Geary, together with their car houses, were converted to serve the large fleet of trolley coaches planned for the rehabilitated system. Each building converted to this use had been a Municipal Railway car house and had been designed with dead-end storage bays, satisfactory for operation of double-end street cars. Both buildings were of reinforced concrete construction; to adapt them for storage of trolley coaches it was necessary to cut an exit opening in the side near the blind end of the bays, remove those adjacent columns which interfered with the resultant right angle turning movement of the coaches, and provide pits and the necessary servicing facilities.

The Potrero yard was completed first

Fig. 2. Trolley coach storage yard (Presidio)

To reduce the complexity of the overhead trolley system, a single pair of contact conductors was provided for every two lines of vehicles; as each coach left the indoor inspection pits and entered the outdoor storage line its poles had to be pulled down so as not to interfere with movement of vehicles in the adjacent line. However, this scheme was found to be undesirable, and when the Presidio yard was designed provision was made for a pair of overhead conductors for each line of vehicles; see Fig. 2. By leaving the trolley poles up at all times considerable saving in hostler's time has been effected and night janitorial work in the coach bodies facilitated.

Coils of 2-conductor neoprene-covered rubber-insulated cable are hung on the walls of the buildings at convenient locations for maneuvering the coaches in relatively inaccessible repair areas. These cables are designed to connect to the ends of the lowered trolley poles of a coach and are energized only during depression of an adjacent push button.

Coaches entering either building for periodic light inspection at the pits must pass over a vertical wire about 18 inches high, resiliently mounted in an insulated floor box. If the resistance between the coach wiring and the coach body has fallen below 5 megohms an alarm sounds when the wire contacts the underframe of the coach. Such coach is immediately taken to an electrical test area. Here both trolley poles are raised into contact with the underside of a copper pan permanently connected to the positive bus of the 600-volt supply system through a 750-volt 750,000-ohm voltmeter, the scale of which is calibrated in megohms. A portable ground cable permanently connected to the negative bus of the 600-volt supply system is then clamped to any convenient metallic part of the coach body. Direct reading of insulation condition can thus be obtained as various contactors are opened and closed, and various conductors in the vehicle are lifted from their terminals in the systematic search for location of the trouble. No hazard from electric shock exists since the maximum current flow is limited by the voltmeter resistance to a harmless 0.8 milliampere.

Before entering storage each trolley shoe must also pass under a contact device such that, if the insert is worn sufficiently, the metal cheeks of the shoe can contact a stationary finger mounted above the trolley conductor. This causes a red lamp to light and an alarm to ring, until turned off manually by the service attendant.

Motor Bus Storage

Two large motor bus yards were designed and built under the rehabilitation program. One, the Ocean Avenue yard, had to be built on two levels because of the topography. This yard is equipped with all the usual facilities for refueling, lubricating, tire and battery servicing, washing, and pit inspection. In addition it is provided with hydraulic lifts, and shop facilities for unit interchange repairs and some light body repairs. It can accommodate from 280 to 300 busses.

The second bus yard, known as Kirkland Yard, can serve and store about 185 busses. Vehicles at this yard in need of anything beyond minor repairs are routed to the central shop at 24th and Utah Streets.

Overhead Trolley System

Construction Program

Installation and rearrangement of trolley contact conductors and feeders was performed under contract in accordance with written specifications and plans prepared by the city's Public Utilities Engineering Bureau. In general, conversion from streetcar to trolley coach operation was effected a line at a time, using gasoline motor busses during the construction period.

An exception to the typical procedure occurred in the case of the Stockton Street line. Here the trolley coach overhead was built alongside or over the existing installation without interruption to streetcar operation. The changeover from streetcar to trolley coach operation was made on a Saturday, using motor busses for one day only during the transition. Noninterfering streetcar conductors and guys still in place after conversion were removed later.

Poles and Eyebolts

Along tangent portions of routes on which streetcars were supplanted by trolley coaches most of the supporting poles for the streetcar overhead system could be reused for the trolley coach system. In some cases, where poles were not otherwise usable because of corrosion at the ground line, they were reinforced in place by addition of a reinforcing ground sleeve.

On new routes, and for the support of switches and special work, some 2,600 tapered tubular steel poles were installed. These ranged in butt size from 8-inch no. 3 gauge single-ply up to 13-inch no. 0 gauge 2-ply. Ground sleeves were not provided.

In the downtown area where the route passed buildings of suitable construction, eyebolts were used in lieu of poles wherever possible. In some section of the outlying districts the Municipal Railway purchased a joint interest in wood poles, owned by others, for support of simple cross spans.

Pole Foundations

All poles were embedded in or bolted to concrete foundations. Design of these foundations was based on the theory that the lateral bearing power of the soil increases directly as depth and that, up to the maximum bearing power, soil resistance is proportional to displacement. Six different sizes of foundation were adopted as standard by the railway to accommodate the entire range of pole sizes used.

Concrete was poured directly into the excavation wherever possible. If forms were required they were removed after the concrete had hardened but in no case less than 4 days after pouring; the space was then back-filled and tamped. Load was not applied to poles until 10 days after pouring.

Conductor Supports

Ordinary supporting spans are of 3/8-inch 7-strand galvanized "Common"-grade guy wire sloped one in five from contact conductor to pole and spaced approximately 100 feet apart. Every other span is an equalizer or feeder, of no. 4/0 gauge 19-strand copper with polyvinyl insulation of 600-volt rating.

Contact conductors on 1-way streets are supported directly on pole brackets. This was justified on the basis of improved appearance, particularly in residential districts, but in addition offers some cost advantage over conventional span construction.

The supporting spans in the Stockton Street tunnel were so short, because of the tunnel arch, that with normal construction they would have been too rigid to permit high-speed vehicle movement. To introduce the desired resilience a heavy coiled-steel spring was inserted at the center of each cross span and pulled up to 500-pound tension. The spans are separated by variable distances (averaging 30 feet) to minimize amplitude of standing waves. Vehicle speed up to 35 miles per hour without danger of dewirement is entirely practical with this construction.

The problem of securing enough resilience to avoid early failure of the contact conductor by fatigue at viaduct underpasses has been successfully solved by

Fig. 3 (left). Resilient bow under viaduct

Fig. 4 (right). Interior of rectifier substation (Twin Peaks)

the device shown in Fig. 3. The flexible linkage supporting the conductors is strung between the two horns of a steel bow of channel section about $5^{1}/_{2}$ feet long. Tension in the linkage, maintained at about 300 pounds, permits the desired vertical movement of the conductors during passage of a coach.

Substations and Feeders

SUBSTATIONS

Before consolidation the Municipal Railway purchased all its energy from the Pacific Gas and Electric Company. This energy, metered at the 600-volt d-c bus, was supplied from eight company-owned substations with a total rated continuous capacity of 18,000 kw. Conversion equipment in these stations is about equally divided between rotary converters and motor generators.

Prior to the consolidation, the Market Street Railway, also purchased its energy from the Pacific Gas and Electric Company. In this case, metering was at the 12-kv bus in each of six substations, five within the city limits and one in Millbrae, a neighboring suburb to the south. Ownership of these substations was divided; the buildings and about half the conversion units (rotary converters) were owned by the railroad; the rest of the conversion units (motor generators) were, and still are, owned by the power company. The total continuous rated capacity of the five substations within the city limits is 31,500 kw (Millbrae station was dismantled after acquisition by the city.)

In July 1945 arrangements were completed by which energy generated by the city's hydroelectric system could be transmitted into San Francisco for use of the various departments and utilities of the city. In the case of the Municipal Railway it is delivered as alternating current at the 12-kv bus of all city-owned substations and as direct current at the 600-volt bus of the eight power company substations.

As streetcar lines were converted to trolley coach and other new coach lines were added under the rehabilitation program it became necessary for the Municipal Railway to construct two additional substations. The first of these, located at the westerly end of Twin Peaks Tunnel, is a 2,000-kw rectifier type. The rectifier, pumps, and switchgear are housed in an attractive, reinforced concrete structure, the north wall of which is formed by the concrete facade of the tunnel entrance. The 12-kv step-down transformer and heat exchanger are located in a fenced enclosure adjacent to the south wall. A small outdoor waiting station is built into a corner niche and forms a part of the substation building. A canopy structure extends over the walkway from the waiting station to an adjacent bus terminal area. The substation, which is unattended, serves four underground feeders entering through ducts into a vault under the floor. Space is provided for future increase in the number of feeders.

The second new substation, also of the unattended rectifier type, is located on the south side of the city near the county line. The rectifier is of sealed tube construction and has a continuous duty rating of 750 kw. Station equipment, in weatherproof cubicles, is erected on a 25- by 56-foot concrete slab surrounded by wire mesh fencing. Two underground feeders from this station reduce load on the nearby Geneva substation during rush hours and permit shutting the latter down during off-peak hours, thus reducing station labor expense.

Although the small Mission substation just described is performing adequately, it is probable that in the future no additional substations smaller than 1,000-kw continuous rating will be installed. It is also unlikely that another outdoor-type station will be built because of the difficulty and hazard of servicing in inclement weather.

FEEDERS

In general, the adequacy of existing feeders and the size and location of additional feeders were determined by spotting the location of vehicles on the system map as they would be at the time of maximum daily vehicle demand. Copper requirements were then determined by the criterion that voltage at the end of any line would not be less than 475 at 50-kw demand per coach. This design, while not the economic optimum, was necessitated by the shortage of copper and the governmental regulations in effect at the time.

With but few exceptions the trolley coach negative system was insulated from the negative rail system except at the substation bus. This insured that the rails, which were often merely covered over in the street rehabilitation program, could be removed later without disturbing the feeder system. It also reduced the electrolysis problems which were still further mitigated by insulating all new underground cable with 600-volt class neoprene-sheathed rubber.

All new underground positive feeder cable was of paper and lead construction, insulated for 2,000 volts and provided with neoprene sheath over the lead. Pole risers were in conduit and the cable was connected and wiped to a pole-top pot head. Because lightning is of rare occurrence in San Francisco no arresters were installed.

Central Control

Important in the rehabilitation program was the establishment of a communication center known as Central Control. The facilities at this center are in a soundproofed room, adjacent to the office of the chief inspector. Two men are normally on duty, a radio operator and a clerk whose duties include watching the headway recorder charts, answering telephones, dispatching repair trucks, and recording delays, accidents, and unusual occurrences.

The base radio station, located in a small concrete building high atop Twin Peaks, is remotely operated from Central Control over a single pair of leased telephone conductors. It is the means of communication with 31 mobile radio transmitters and receivers, installed in inspectors' automobiles and on tower trucks and emergency vehicles. A fully automatic gasoline-electric power plant is maintained at the base station, and another at Central Control to prevent shutdown of the radio in event of power failure.

The headway recorder charts provide a record of the times at which inbound and outbound vehicles pass by 40 different points on the system. The instruments are under constant observation and in the event of an abnormal delay in vehicle passage, an inspector is dispatched by radio to investigate. At the end of the day the charts are sent to the schedule department as an aid to realistic schedule construction and modification.

No Discussion

An Extension of the Root Locus Method to Obtain Closed-Loop Frequency Response of Feedback Control Systems

ALBERT S. JACKSON
ASSOCIATE MEMBER AIEE

THE root locus method for analyzing feedback control systems has become increasingly popular over the last few years. However, the great potentialities of the method have yet to be realized. The method has had the drawback of presenting only the transient response of the system under study. An ideal method of analysis and synthesis would allow the engineer to arrive at both the transient and the steady-state sinusoidal response without any large amount of extra work involved for either. The work would progress in two stages. The first stage would take a relatively short time to complete and would yield approximations to both the closed-loop transient response and the frequency response. At this point, various design features could be compared. The second stage would be an extension of the work done in the first stage—not a more detailed rework of the problem. This second stage of the problem would yield exact and complete transient and sinusoidal solutions.

The root locus method performs this function quite well with regard to the transient response. In this paper an extension of the method is presented which leads to the closed-loop sinusoidal response at the same time that the transient response is obtained. The functions of a near ideal method of analysis and synthesis are thereby performed by the root locus method.

Root Locus Method

As the name implies, root locus is a method whereby the loci of the roots of the characteristic equation are plotted with the open-loop gain as the parameter. Consider the linear system shown in Fig. 1, where

$G(s)=K_1\dfrac{G_1(s)}{G_2(s)}$, Laplace transform of forward transfer function

$H(s)=K_2\dfrac{H_1(s)}{H_2(s)}$ Laplace transform of feedback or reverse transfer function

$K=K_1K_2$, Open-loop gain of system

Then the closed-loop response is as follows

$$\frac{C}{R}(s)=\frac{G(s)}{1+G(s)H(s)}=\frac{K_1G_1(s)H_2(s)}{KG_1(s)H_1(s)+G_2(s)H_2(s)} \qquad (1)$$

The transient response is determined by the roots of the characteristic equation

$$1+G(s)H(s)=0 \qquad (2)$$

These roots determine the stability and speed of response of the closed-loop system. Therefore, these roots are of primary interest.

Root locus is a method of graphically solving for these roots of the characteristic equation. This is accomplished by setting $G(s)H(s)=-1$. Since this is a complex equation, the two relations

$$|G(s)H(s)|=1 \qquad (3)$$

$$\arg G(s)H(s)=\pm n\pi \quad (n=\text{an odd integer}) \qquad (4)$$

are obtained. Equation 4 can be used to plot the loci of the roots in the s-plane and equation 3 used to determine K at points along these loci. The actual mechanics of plotting the loci are quite simple. One method is to find points by trial and error in the s-plane such that the sum of the arguments of all the complex vectors from the various poles and zeros of $KG(s)\ H(s)$ to the point chosen is an odd multiple of π. The open-loop gain K corresponding to a particular point of the locus is found by measuring the distances from the poles and zeros to the point in question and performing the operation of equation 3. The actual mechanics involved can be performed by use of a Spirule,[1] special plotting boards, electro-

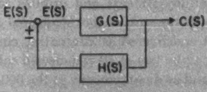


Fig. 1. Block diagram of feedback control system

Paper 54-281, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted February 10, 1954; made available for printing May 12, 1954.

ALBERT S. JACKSON is with Cornell University, Ithaca, N. Y.

The author acknowledges the helpful suggestions and co-operation of Prof. W. E. Meserve of Cornell University.

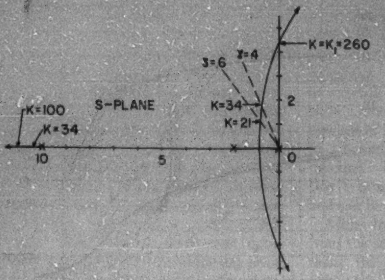


Fig. 2 (left). A typical root locus plot

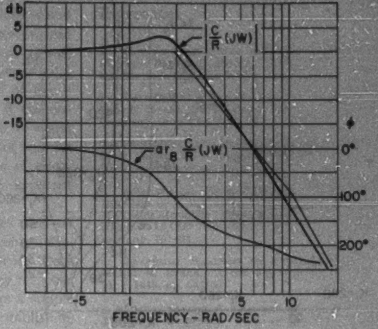


Fig. 3 (right). The resulting frequency response curve for K=34 from Fig. 2

lytic tanks, or simply a protractor and pair of dividers. Fig. 2 shows a typical locus of roots in the s-plane.

It is seen then that the loci of the roots of $KG_1(s)H_1(s)+G_2(s)H_2(s)=0$ are obtained quite readily for all values of K. Thus the transient response of the closed-loop system may be obtained for a step input. If the initial conditions were known, the complete time response of the system to any input could be evaluated if it were desired.

Frequency Response

It is desirable to use this same information, i.e., the root loci, to determine the sinusoidal response of the control system. The steady-state response of the closed-loop system, as described by equation 1, to a sinusoidal input is

$$\frac{C(t)}{Fe^{j\omega_t t}}=\frac{C}{R}(j\omega_t)$$
$$=\frac{K_1G_1(j\omega_t)H_2(j\omega_t)}{KG_1(j\omega_t)H_1(j\omega_t)+G_2(j\omega_t)H_2(j\omega_t)} \quad (5)$$

This is the familiar form employed to obtain the frequency response of a transfer function. That is, if one knows $(C/R)(s)$, the frequency response may be obtained by substituting $s=j\omega$ and letting ω take on all values from zero to infinity. If the factors of both numerator and denominator are known, the method of Bode[2] can be used to obtain the log amplitude and angle versus log frequency curves. Now for any particular value of open-loop gain, the factors of the denominator are the roots of the characteristic equation and are found from the root loci. The factors of the numerator are the zeros of the forward transfer function and the poles of the feedback transfer function. These zeros of $G(s)$ and poles of $H(s)$ had to be known to plot the root loci and are found on the s-plane plot. Therefore, all of the factors of $(C/R)(s)$ are found from the root locus plot and the method of Bode can be used to plot very rapidly the closed-loop frequency response of the system under study. Suppose the root locus plot for a particular system is that of Fig. 2. It is seen that the system is unstable for $K_1=260$. The damping of the complex conjugate roots can be brought to a reasonable value by reducing the open-loop gain to a range of 20 to 35. A gain of 34 would give a per-unit critical damping of $\zeta=0.4$. For this condition, the resonant frequency of these underdamped roots is 1.7 radians per second and the natural undamped frequency 1.8 radians per second. The other root is located at $S=-10.5$. Thus the closed-loop Laplace transform is

$$\frac{C}{R}(s)=\frac{34}{(s+10.5)(s^2+1.52s+1.8^2)}$$

The resulting log frequency plot is shown in Fig. 3. There is a quadratic break point at $\omega=1.8$ radians per second and a single break point at $\omega=10.5$ radians per second. The per-unit critical damping ζ associated with the quadratic break point is 0.4. The exact curves are obtained by adding corrections to the straight line asymptotes. These corrections can be easily calculated or taken from prepared charts.[3]

From equation 5 it is seen that the frequency response could also be obtained by determining in the s-plane all of the factors $(j\omega_t-r_k)$ point by point for each value of ω_t from zero to infinity and calculating $(C/R)(j\omega)$ from equation 5. The same result would be obtained by writing equation 1 in factored form, the roots of the denominator being obtained from the root loci, and calculating the frequency response by substituting in $S=j\omega_t$. The Bode method accomplishes the same thing with much more ease.

It is interesting to note that a zero in $H(s)$ or the same zero in $G(s)$ will yield the same root loci and therefore the form of the transient response will be the same. However, the zero in $H(s)$ has no direct effect upon the frequency response, whereas the zero in $G(s)$ contributes a factor directly to the frequency response. (The reverse is true for poles.) This is demonstrated in Fig. 4. If the zero in question is a factor of $G(s)$ as in Fig. 4(B), a leading break point appears in the frequency response plot which was not present in the case of Fig. 4(A). Note that for a particular value of K all of the other break points remain the same.

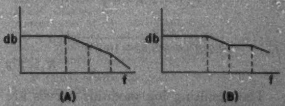
Illustrative Example

Consider the simple type-1 system described by

$$G(s)=\frac{K_1}{S(s+2)(s+10)}$$

$$H(s)=1$$

The root loci of this system are shown in Fig. 2. It is desired to pull the loci of the two complex conjugate roots back over into the left half of the plane to insure stability and, at the same time, to improve the bandwidth and transient response. It is also desired to improve the velocity error constant if possible. The most common way of stabilizing such a system is by use of tachometer feedback. This is called tachometer damping or is sometimes referred to as proportional plus rate control.



(A) (B)

Fig. 4. Frequency response asymptote for systems having the same G(s)H(s) but with zero supplied (A) in H(s) and (B) in G(s)

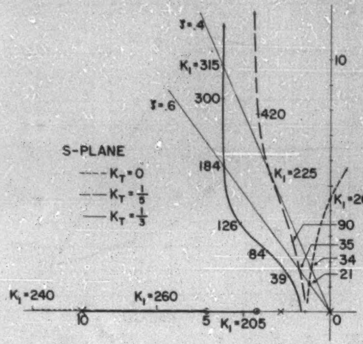


Fig. 5 (left). Root locus plots for three values of K_t

$$G(s)=\frac{K_1}{s(s+2)(s+10)}$$

$$H(s)=1+K_tS$$

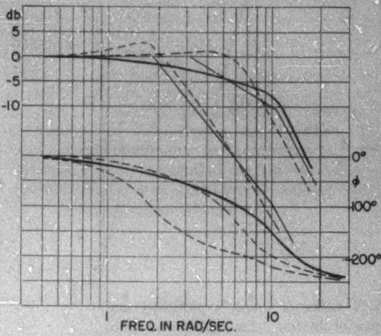


Fig. 6 (right). The resulting frequency response curves from Fig. 5 for $\zeta=0.4$ and

- - - - $K_t=0$
— — $K_t=1/5$
—— $K_t=1/3$

Let K_t be the gain factor of the tachometer; then $H(s)=1+K_tS$. This introduces a zero at $S=-1/K_t$ in the s-plane. The resulting root loci are shown in Fig. 5 for three tachometer gain settings: $K_t=0$ 1/5, and 1/3. The forward loop gain K_1 is shown as the loci parameter.

It is seen that tachometer damping has the effect of pulling the loci of the conjugate roots over into the left half of the s-plane for all values of gain. Furthermore, it changes the gain scale and completely alters the locus of the real root.

Suppose the designer is interested in a transient response which has a value of $\zeta=0.4$ for the per-unit critical damping for the least damped roots. Using $K_t=1/5$, the forward loop gain can be boosted from 34 to 255 and still have $\zeta=0.4$. For $K_t=1/3$, K_1 can be increased to 375. Although ζ has remained the same, the rise time due to the underdamped roots has decreased considerably and ω_r has increased. However, the position of the real root, which helps determine the total rise time, has changed for the worse. For $K_t=1/3$, the time constant of this real root has become larger than that of the underdamped roots.

To see what all of this means in terms of the frequency response, the closed-loop transfer functions as taken off the root locus plot for $\zeta=0.4$ are

$$\frac{C}{R}(s)=\frac{34}{(s+10.5)(s^2+1.5s+1.8^2)}\text{ when }K_t=0$$

$$\frac{C}{R}(s)=\frac{255}{(s+7.1)(s^2+4.5s+6^2)}\text{ when }K_t=1/5$$

$$\frac{C}{R}(s)=\frac{375}{(s+3.2)(s^2+8.6s+10.8^2)}\text{ when }K_t=1/3$$

The resulting log modulus curves are plotted in Fig. 6. The original system had a cutoff frequency of 2.6 radians. By using $K_t=1/5$, this is extended to 7.0 radians. However, for the larger amount of tachometer damping of $K_t=1/3$, the out-off frequency has been reduced back to 3.8 radians. Added tachometer damping is seen to move the quadratic break point out but, at the same time, to move the single break point in. The widest bandwidth for $\zeta=0.4$ is obtained by using a slightly larger value of K_t than 1/5.

Attention should be given to the resulting phase shift of the closed-loop system. The phase shift is improved considerably for both cases of tachometer damping. The drooping amplitude curve for $K_t=1/3$ gives the least phase shift at frequencies above 2 radians per second. The phase shift could be improved by moving the quadratic break point out, but this results in lowering the frequency of the single break point as seen from the root locus plot of Fig. 5.

The value of $\zeta=0.4$ for the per-unit critical damping of the complex roots was chosen quite arbitrarily for the foregoing example. As it turns out, the frequency response for $K_t=1/5$ is quite good. However, if the single break point were moved in to about 6 radians per second and the quadratic break point moved out to 7 or 8 radians per second, the bandwidth could be increased even more. By inspection of the root locus plot of Fig. 5 it is seen that a gain of approximately 400 would give a quadratic break point at 8 and a single break point near 6.2 radians per second. This would extend the bandwidth to about 10 radians per second. The single break point would smooth out the resonant peak, resulting in a very smooth frequency response, even though the damping of the complex roots has decreased.

Even for this simple example, it would have taken considerable time to obtain the frequency response curves of Fig. 6 by the usual methods using the Nichols' chart or the Nyquist plot. In fact, it would have taken at least as much time as was required to obtain the entire results of both Figs. 5 and 6, from which complete transient and frequency response behavior can be obtained. Furthermore, the gains K_1 and K_2 would normally be chosen by using the M-criterion where

$$Me^{j\alpha}=\frac{G(j\omega)}{1+G(j\omega)}$$

This assumes unity feedback, although the criterion can be modified to account for real, frequency invariant values of $H(s)$ other than 1.[4] The method is further restricted and often misused by the fact that the M-criterion is related to the transient response only in the case of a quadratic system, i.e., one having a quadratic characteristic equation. The possible pitfalls of using the M-criterion for higher order systems are pointed out quite effectively in the example problem. Consider the log modulus plot for $K_t=1/5$ in Fig. 6. The maximum value of $(C/R)(j\omega)$ is +1 decibel or $M_m=1.12$. This should correspond to a transient response wherein the least damped roots have a per-unit critical damping of $\zeta=0.53$. From the root locus plot, however, it is known that the actual value is $\zeta=0.4$. The M-criterion applies even less to the case with $K_t=1/3$, for there the maximum value of $(C/R)(j\omega)$ is unity, which would correspond to a critically damped or overdamped system.

Conclusions

The root locus method has been considered to be strictly a transient response method, as opposed to the frequency response methods of Nyquist and Bode. Actually, the root locus method gives information which enables one to calculate the response to any input. In particular, the steady-state response to a sinusoidal input is obtained with very little additional work. By using Bode's method, the closed-loop frequency response can be obtained very quickly from the information contained in the root locus plot. Thus the powerful concepts of the frequency response methods may be combined with those of the transient response methods when the root locus plot is used to analyze or synthesize a control system. The closed-loop transfer function need not be of the minimum-phase type for the Bode method to apply. However, it must be known whether the particular transfer function is or is not of the minimum-phase type in order to apply it. The root locus method has a decided advantage over other methods here since all of the factors, both zeros and poles, are known for all values of gain once the loci have been located. Thus, suppose a pole of $H(s)$ or a zero of $G(s)$ were located in the right half of the s-plane. For a certain range of K the system could be quite stable, but it would be a nonminimum-phase type of transfer function. Using the concepts of equation 1 this situation is readily apparent and can be taken into account.

Aside from the advantages derived from the study of particular control systems, the extension of the root locus method to obtain the frequency response of the closed-loop system offers a simple and very effective means of establishing correlation between the transient response and frequency response of feedback control systems.

References

1. Control System Synthesis by Root Locus Method, Walter R. Evans. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 66–69.

2. Network Analysis and Feedback Amplifier Design (book), H. W. Bode. D. Van Nostrand Company, Inc., New York, N. Y., 1945.

3. Principles of Servomechanisms (book), G. S. Brown, D. P. Campbell. John Wiley and Sons, Inc., New York, N. Y., 1948.

4. A Graphical Procedure for Determining the Gain of a Servomechanism for a Specified Maximum Modulus Less Than Unity, Thomas J. Higgins. *AIEE Transactions*, vol. 73, pt. II, July 1954, pp. 101–04.

5. Graphical Analysis of Control Systems, Walter R. Evans. *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 547–51.

6. Servomechanism Analysis (book), G. J. Thaler, R. G. Brown. McGraw-Hill Book Company, Inc., New York, N. Y., 1953.

7. Transients in Linear Systems (book), M. F. Gardner, J. L. Barnes. John Wiley and Sons, Inc., New York, N. Y., 1942.

No Discussion

Performance of a Constant-Speed Drive

E. W. GILOY
ASSOCIATE MEMBER AIEE

Synopsis: Aircraft application of regulated frequency a-c power obtained from main-engine driven sources requires a generator drive capable of providing a controlled output speed independent of airplane engine operation. Performance of a variable ratio hydraulic-mechanical drive utilizing engine shaft power for this purpose is presented. General characteristics and service experience covering an actual installation are included.

THE application of alternating current to aircraft as the basic electric-power type will in general involve a requirement for frequency control. Frequency wild systems, while capable of powering resistive type loads, do not adequately meet the requirements of large present-day and future aircraft. It is improbable that a basic a-c airplane would exist if it had been necessary to accept the restrictions of variable-frequency a-c power. However, the feasibility of a controlled frequency a-c system powered by an aircraft's main propulsion units is not accepted by all aircraft engineers. Aircraft propulsion engines are essentially variable speed units operated in accordance with over-all airplane requirements which are generally directly opposed to generator drive considerations. An objective, therefore, exists for an engine power take-off capable of providing a generator drive speed which may be controlled independently of engine operation. To date various methods and devices, generally classed as constant-speed drives, are in development or in limited use to perform this function. One such device of the hydraulic-mechanical type utilizing engine shaft power has recently been applied to a present-day patrol type of airplane (see Fig. 1) and it is the purpose of this paper to present the operation and performance obtained with this installation to date. The application is believed to be the second one of its general type involving an in-production airplane. Approximately 6,000 hours of constant-speed drive operation have been logged as of this writing.

Electric System

The electric system associated with the constant-speed drive application is a 3-phase wye-connected 3-wire grounded neutral 115/200-volt 400-cycle a-c system. Airplane requirements for regulated frequency power include the flight control, electronic, defense, cabin conditioning, engine fuel control and hydraulic systems. Electric power sources are three 40-kva 30-kw 6,000-rpm a-c generators. Two generators are indirectly main-engine driven by constant speed drives (one per engine) and the third machine is directly driven by a regulated gas turbine type of auxiliary power unit. The three a-c generators are connected to a sectionalized bus which may be electrically connected as a single unit to allow for parallel operation of main-engine generators. The auxiliary power unit a-c generator, although parallel operation was desired, cannot be paralleled with the main-engine a-c sources. The auxiliary power unit is not equipped with a governor system and control suitable for parallel operation, and problems associated with unit procurement prevented the necessary modifications.

Airplane d-c requirements are supplied by a 28-volt d-c system consisting powerwise of two 6-kw 30-volt main-engine driven generators, a single 34 ampere-hour battery and a 6-kw regulated transformer-rectifier converter operating from the a-c system.

Constant-Speed Drive System

The constant-speed drives are the close-coupled differential action type of variable ratio hydraulic-mechanical transmis-

---

Paper 54-217, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted January 8, 1954; made available for printing April 12, 1954.

E. W. Giloy is with the Glenn L. Martin Company, Baltimore, Md.

The author takes this opportunity to thank the Sundstrand Machine Tool Company and members of The Glenn L. Martin Company Engineering Division for their co-operation in the preparation of this paper.

Fig. 1. Patrol-type airplane with constant-speed drive device

sion.[1] The transmission oil system is independent of the airplane engine oil system and engine shaft power is received from the engine accessory pad through a pad adaptor and a short length of flexible shaft using two Hooke's joints and a spline. Installation of the drive transmission in the nacelle aft of the fire wall is shown in Fig. 2. The drive transmission (see Fig. 3) is complete with free-wheeling clutch, governor and control assembly, overspeed solenoid valve, under- and overspeed switches, and charge and scavenge pumps. An accessory electrical control box is provided for trimming control of transmission output speed and to control load sharing between paralleled units. Action of this electrical control will be described.

It is probable that the reader is at least generally familiar with the method of drive transmission operation. A detailed explanation, therefore, is not included. Suffice it to say that transmission of power through the unit is accomplished by mechanical as well as by hydraulic means and that the unit consists essentially of a variable-displacement axial piston-type pump and a fixed displacement axial piston-type motor. Pump volumetric displacement is varied in relation to the transmission input speed to provide the essentially constant-speed output.

The weight breakdown of a complete single-drive system installation is given in Table I.

Governing Method

The method of drive governing used is different from that previously used for this type of generator drive. Governing is fundamentally provided at the drive transmission and is basically independent of the electrical control used for close regulation of drive output speed and torque. The governor is a mechanical type of hydraulic control valve-spring biased and fly-ball operated. Speed sense is obtained from the output of the transmission ahead of the free-wheeling clutch. The governor controls the porting of regulated control pressure oil to the hydraulic servo actuating the wobbler plate which regulates the displacement of the pump end of the transmission. For trimming control of drive output speed, a means is provided to change the total spring biasing of the governor by the action of a small secondary spring located inside the main governor spring. A plunger rod rests on the upper free end of this secondary spring and rides against a cam plate driven by a small 2-phase a-c motor. Operation of the motor and the resulting repositioning of the cam provides the desired small change in the total spring reference of the governor and thereby a change in the regulation point of the governor.

Control for the governor a-c trimmer motor is the magnetic-amplifier type and is provided by the drive control box using a voltage and current signal obtained from the generator being driven. Magnetic amplifier action is regulated by a frequency discriminator circuit as well as a load-equalizing control circuit. The load control circuit is used only during parallel operation. Manual monitoring of drive speed and load sharing is provided at the airplane electric power control panel by a potentiometer connected across the output of the drive frequency discriminator circuit. During parallel operation, the load control circuits of the drives are interconnected to effect load equalization while maintaining system frequency.

Table I. Single-Drive System Installation

Parts	Weight, Pounds
Transmission (complete)	76.0
Control box	9.5
Current transformer	0.5
Potentiometer	0.3
Wiring	4.1
Air filter	0.25
Control box installation	0.7
Engine pad adaptor	7.5
Flexible shaft	6.84
Oil cooler	6.75
Oil lines	4.8
Oil filter	2.5
Oil reservoir and vent	5.2
Fire wall seal and shroud	4.01
Oil cooler air ducts	7.57
Transmission installation (includes generator provisions)	12.7
Oil cooler installation	1.82
Oil reservoir installation	1.0
Oil transmission	2.9
Oil lines	1.34
Oil cooler	3.25
Oil reservoir	13.22
	172.75

The variation of drive speed that can be obtained by use of the aforementioned potentiometer and the associated governor trimmer system is intentionally limited by the design and construction of the cam driven by the governor trimmer motor. The limits of generator frequency that can be obtained by full travel of the cam are 385±5 to 415±5 cycles per second (cps). This feature of drive governoring insures that no action or lack of action of the governor trimmer system can result in a generator frequency (drive output speed) outside the permissible operating limits of the airplane's electrical and electronic equipment.

It is believed that this combination of mechanical flyball governor and electrical trimmer system offers an increase in overall reliability and invulnerability over the straight electromechanical type of governor. With the mechanical flyball governor system there is no extension of basic governor control beyond the confines of the unit being controlled. Damage to the extended trimmer system will affect parallel operation of the drives

Fig. 2 (left). Installation of drive transmission in the nacelle aft of the fire wall

Fig. 3 (right). Complete drive transmission

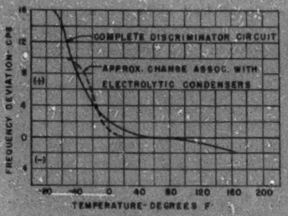


Fig. 4. Frequency discriminator temperature characteristics. Frequency initially adjusted to 400 cps at 76° F

however, single drive operation under this condition is quite feasible. The sectionalized bus design of the airplane allows for this contingency.

Drive Characteristics

The input speed range of the drive for governed operation is 2,600 to 9,500 rpm. Generator loads up to 15 kw can be carried with drive input speeds below 3,000 rpm. Above 3,000 rpm, allowable generator loading can be increased to 30 kw for short-time operation. For this application, however, engine pad limitations raise this speed requirement to 3,300 rpm. At 4,350 rpm and higher, the full design capabilities of the drive are available, and they provide a generator output of 30 kw continuously and overloads of 45 and 60 kw for 5 minutes and 5 seconds respectively. The design base speed of 4,350 rpm for the drive provides for its full use over all airplane flight conditions including engine minimum cruise rpm to military power conditions. The minimum governing input speed of 2,600 rpm allows drive use for engine-idle conditions down to approximately 800 engine rpm. Below a drive input speed of 2,600 rpm the transmission operates as a fixed ratio drive with a resulting decrease in generator frequency with decreasing input speed. At a generator frequency of approximately 365 cps the generator is disconnected from the electrical system by action of the drive underspeed switch. In the event of an overspeed condition, i.e., exceeding 7,000 rpm, the drive is put in full underdrive condition by operation of the drive overspeed switch.

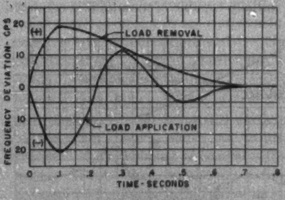


Fig. 5. Frequency transient characteristic for full-load shock condition. Drive input speed 7,100 rpm no load, 7,000 rpm full load

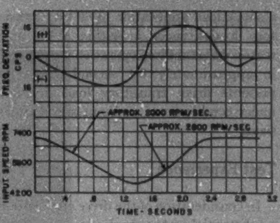


Fig. 6. Frequency transient characteristic for input speed change. Generator at no load

Speed Regulation

Drive speed regulation under stabilized conditions is within ±1/2 per cent of the selected speed setting (normally 6,000 rpm to provide a generator frequency of 400 cps) for all conditions of design load, overload, and input speed. This degree of speed regulation is obtained by action of the frequency discriminator control. The basic flyball governor characteristic, unmonitored by the trimming control, would result in a 2- or 3-cycle droop during warm-up as well as a 4-cycle droop under full-load conditions. Basic governor droop under a 150 per cent load condition would be approximately 10 cycles.

The frequency discriminator circuit used for close regulation of drive output speed is not completely independent of temperature effects although the change in speed setting with temperature is within ±2 cycles of normal generator frequency over a rather large temperature range. Fig. 4 shows the effect of temperature on the frequency discriminator circuit and indicates that with a drive speed originally set for a generator frequency of 400 cps at a temperature of 76 degrees Fahrenheit (F), a decrease in temperature to 0 degrees F results in a frequency increase of two cps. From 0 degrees F to −65 degrees the frequency rises rapidly. For an increase in temperature the frequency change is much less severe and does not exceed 2 cps up to 165 degrees F. The change in frequency setting with temperature is chargeable to a change in capacity of the condensers associated with the discriminator circuit and as further shown by Fig. 4 results in large part from the capacity change of the electrolytic condensers used to filter the output of the discriminator.

Fig. 5 presents a representative curve of drive transient response for full-load shock conditions, input speed to the drive being held nearly constant and full-load shock applied from a no-load generator condition and then shock removed. Drive over-all transient response time is shown to be approximately 0.7 second and the maximum instantaneous change in generator frequency is approximately 20 cps as an average for this condition. A maximum-frequency deviation of from approximately 15 to 30 cps can be obtained, however, for this test depending on the drive wear condition. A similar characteristic will be obtained with the governor trimmer system inactive, indicating that the response shown is largely independent of the trimmer system. The maximum transient frequency deviation is reasonably linear for other magnitudes of shock load.

Drive transient response under conditions of an input speed change is shown in Fig. 6. The condition presented involves a rather large acceleration and deceleration rate resulting in a maximum generator frequency change of approximately 16 cps. For an input speed change of 1,000 rpm per second the maximum frequency change is approximately 4 cps. At an acceleration or deceleration rate of 500 rpm per second, frequency change is approximately 1 cps. Fig. 7 presents a plot of this characteristic.

Efficiency

Drive over-all efficiency including transmission charge and scavenge pump losses varies from approximately 87 to 73 per cent. The efficiency of the drive is a function of input speed and is maximum at approximately 4,000 input rpm. Drive straight-through operation occurs at approximately 6,600 input rpm but maximum efficiency is not obtained under this condition because of the pumping losses

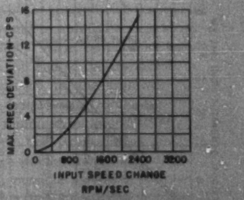


Fig. 7. Maximum transient frequency deviation with input speed change

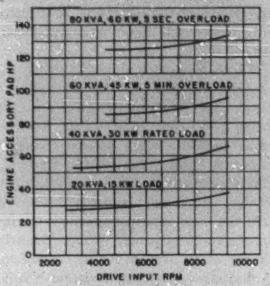


Fig. 8. Drive-generator horsepower requirements

of the charge and scavenge pumps. The horsepower requirements for a complete generator and drive installation as referred to the airplane engine accessory pad is given in Fig. 8.

Parallel Operation

Real load sharing between generator-drive combinations is in general within 3 kw for all stabilized conditions of generator load and drive input speed. During large overload conditions, however, load unbalance may exceed this figure and, as indicated by Fig. 9, may be approximately 10 kw under conditions of approximately 200-per-cent load. Under no-load conditions one generator-drive combination may tend to motor the other unit. Generator input to overrun the drive clutch is about 2 kw and this magnitude of load is marginal for the sensitivity of the load division control. More information relative to load division sensitivity is given later.

Under shock load conditions drive parallel operation is stable and transitory load division is good. For example, for a shock load application of 60 kw, transitory load division is within 4 kw and reasonably stabilized operation is obtained in approximately 0.1 second. Fig. 10 presents drive transitory load division under conditions of an input speed transient.

Test Experience

In consideration of the drive application a substantial test evaluation was made of drive performance. The program was threefold, involving laboratory and airplane evaluation at The Glenn L. Martin Company's plant, and laboratory testing with the aid of the facilities of the drive manufacturer. Laboratory work at the Martin Company was associated with a setup of the airplane's electrical system (Figs. 11 and 12) and provided information relative to over-all system operation. The program included single and parallel operation and demonstrated to a large extent the over-all practicability and reliability of the unit. Testing was at times severe, and peak short-time loads as large as approximately 95-kw generator load were recorded for a single drive. The drive on which most of this overload and shock testing was done logged approximately 250 hours of test operation during this phase of the program.

The complete test and evaluation program was not without incident, however, and items of difficulty were encountered. Space and, perhaps, the over-all value to the reader do not warrant a complete treatment of the subject many of the items being relatively minor and involving a finalization of the development and a "shakedown" of the new governor system. There were several items, however, which it is believed will interest the reader and these experiences are given in the following.

Perhaps the most basic item noted early in the program concerned drive load stability in parallel operation. It was found that a condition could exist in which the drives would shift the entire load from one to the other in an oscillatory manner at a rate approximating a cycle of load shift per second. The condition was marginal and it varied with drive input speed, freeness of the governor trimmer motor system, and other items affecting over-all drive sensitivity. The slow rate of the oscillatory condition made electrical stabilization of the governor-trimmer motor-control circuit difficult. Mechanical damping was considered but was not believed feasible in consideration of environmental requirements and the accuracy of damping required. Control of the condition has been obtained, however, by limiting the sensitivity of the drive load division circuit and by introducing a degree of nonlinearity in its response through the use of shunt-connected germanium diodes across the magnetic amplifier control winding. The resulting over-all sensitivity and strength of the drive load division control is generally effective although occasional manual monitoring of drive load division control is required. Sensitivity is low and requires, in the present design configuration, very low voltage performance for the governor-trimmer motor and associated assembly. For satisfactory load division the motor must respond to control-winding voltages as low as 5 volts. Available motor torque at this condition requires an optimum operating assembly of motor, motor speed reduction gear box, cam follower, etc. It is believed desirable for future designs of this type that the sensitivity and strength of load division control be increased and in particular that a higher allowable minimum control-winding voltage for motor operation be permissible. A stronger motor torque effort is required to provide a governor trimmer assembly less subject to the variations of manufacture, wear, etc. Improvement or schemes to improve drive load stability during parallel operation will undoubtedly be required to achieve this. As an aid to drive stability, the incorporation of mechanical lead in the basic hydraulic governor is desirable.

During overload testing of the drive it was observed that the drive was not initially capable of maintaining regulated

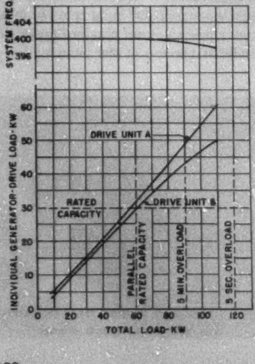


Fig. 9 (left). Load division characteristic. Stabilized condition

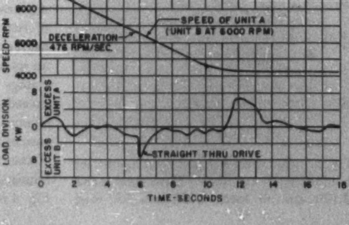


Fig. 10 (right). Load division during input speed transient. 50-kw total load

Fig. 11 (left). System test setup showing control panel and bus center

Fig. 12 (right). A-c and d-c generator test setup

output speed under conditions of large overload. At approximately 55-kw generator load, the servo holding the pump wobble plate would yield and generator frequency would drop to something under 350 cps. Servo positioning pressure is obtained from the charge oil system and an investigation of the problem indicated that during large overloads, transmission slip losses as well as other leakage losses were resulting in a reduction of charge oil pressure. This fact accounted for the loss of drive speed regulation. Charge pressure is nominally adjusted by relief valve setting to 400 pounds per square inch (psi). Under 100-per-cent overload conditions charge pressure was dropping to approximately 250 psi. Governor pressure required under this load condition to hold the pump wobbler is approximately 305 psi. Corrective action for this item produced what has been called the "blow by system." With the blow by system transmission slip oil is returned directly to the charge oil section and not to the transmission sump as previously. This action maintains charge oil pressure during overload conditions and thereby provides sufficient governor pressure for drive regulation under overload.

Airplane evaluation brought to light a condition of drive operation that did not present itself during any phase of laboratory testing, at least not to a degree which was recognized. The condition might be called rough governing or an erratic modulation of the drive output speed. The condition is not present in all drive installations and is not of a fixed intensity where it does occur. Where present the phenomenon, as observed on a visual type of frequency meter, shows an erratic (nonoscillatory) bobbing of generator frequency about the regulating point. Average regulation is within drive limits but frequency has been observed to jump 3 or 4 cycles on an instantaneous basis. Qualitative analysis of the condition has not been made, and therefore the extent or rate of the drive speed modulation cannot be stated with accuracy. During parallel operation, the steadiness of regulation is improved because of the nature of such operation, but load division between drives is jumpy. Tests on the drive with the pump wobbler blocked indicate that the condition is the result of engine torsional vibration transmitted through the drive transmission. This fact would account for its absence during laboratory testing in which an electric motor drive was used. Perhaps a torsional damper in the transmission input shaft is required to eliminate the trouble. Along this line, in an effort to improve the steadiness of drive governing, it is planned to investigate and possibly to use a governor drive that does include some degree of torsional damping. It is too early to indicate the success of this move. No trouble to the drive or to the airplane's electric system is known to exist because of this condition but solution of the problem is believed desirable if a proper standard of quality for aircraft a-c systems is to be obtained.

Service Experience

As of this writing the drive has a total of approximately 6,000 hours of service operation and it is understood that the drive and associated electric system have been well received by operating personnel.

Drive troubles in service have largely involved items of the nuisance type associated with the shakedown phase usually experienced in a new application. For example, drives have been damaged owing to oil line failures. This item has been corrected by an improved technique of oil line manufacture. Trouble was experienced with operating personnel standing on the drive governor assembly, particularly the trimmer motor. A step-type guard was installed to protect this assembly. Failures of a bearing associated with the drive over and under speed switches were experienced, and corrective action on this item is presently in progress. Other troubles include single cases of a transmission input shaft failure, an output shaft seal leak and one transmission damaged by a break-up of an adjacent piece of equipment. Drive replacement necessitated by problems of the type noted has approximated 10 per cent but is decreasing as corrective action for the various items is taken.

Superimposed on this normal pattern of shakedown in service, there has been a recent series of serious malfunctions and internal breakups involving a number of drives, although extensive flight hours at The Glenn L. Martin Company, as well as customer airplane test programs and approximately 4,000 hours of initial service operation, did not indicate trouble of this type. This new problem appears to involve specific engine operating techniques affecting the drive blow by system and producing separation. For normal and proper operation of the drive, the respective pump and motor pistons and push rods must be in firm contact with each other and with their respective wobbler faces. Any separation between these parts of the transmission during operation can result in serious damage to the transmission. With the blow by system, the push rod end of the transmission pistons is exposed to the transmission charge pressure. During zero or light load there is a minimum of pressure difference between the transmission-working and charge pressures, and separation is most often possible at this time if the transmission should be subject to a shock condition involving a large load dump or a rapid input speed change where inertial effects may tend to initiate some degree of separation. Therefore, indications are that the problem is associated with engine backfiring and the attendant very rapid decelerations and accelerations of drive transmission input speed. Engine backfires are generally incurred during starting and at low idle conditions where transmission pump wobbler angles are large and the overhang of the pump push rods is at a maximum. At this time the mechanical effects of separation are most severe. It is not known at this moment what the action on this item will be. Perhaps removal of the blow-by feature from the pump end of the transmission may relieve this trouble and still retain desired drive overload regulation. It also may be possible to reduce the transmission-pump wobbler moment under overload conditions to a value within the capabilities of the governor pressure without the blow-by feature. Complete evaluation of the problem is necessary, however, before the proper move can be ascertained. Drive replacement, as influenced by this latest item, has reversed its previous downward trend and over-all replacement now amounts to approximately 20 per cent of the drives in service.

At this writing, accumulated service time for individual drives varies from a low of approximately 12 hours to a high of 800 hours with the average accumulated time at the moment being approximately 120 hours. Two units have reached the 800-hour figure and were removed for overhaul at that time. These drives were in good operating condition but removal was advised because of wear on the worm drive for the governor.

Conclusions

The over-all operation and performance of the constant speed drive and associated electric system have been gratifying. As a result of the system, airplane operation has incurred practically a minimum of delay, at least up until the more recent problem involving drive operation. This problem is certainly temporary, however, and implies no basic fallacy as to the feasibility and practicability of this type of generator drive utilizing engine shaft power. All types of generator drives must meet the requirements of actual service environment and this experience is merely another step toward that end.

For future designs and applications of this type of drive system and, for that matter, other types of drive systems, it is considered that drive improvement should be directed along these general lines. Use of a basic mechanical type of governor coupled with an electrical means for trimming control is believed sound, and further development of its design should be encouraged. Load division control needs strengthening with the objective of eliminating the need or use of any remote manual adjustment. The accuracy of frequency regulation over all environmental conditions should be improved and attention should be given to the instantaneous as well as the average value of regulation. Efforts to increase the response of governing and thereby minimize transient conditions are also desirable. Greater effort to design for actual airplane operating conditions is necessary and a proper union of generator and drive capabilities must always be considered.

It is hoped that this detailed and frank discussion of laboratory and service experience will contribute to a more general understanding of the basic reliability and performance now realizable with the regulated frequency a-c system for aircraft; serve to indicate the areas where further effort should be extended to improve reliability and performance standards.

References

1. A CONSTANT-SPEED DRIVE FOR AIRCRAFT ALTERNATORS, L. H. Schuette, R. Chrzanowski. *Applied Hydraulics*, July 1948.

2. CONSTANT-SPEED DRIVES FOR AIRCRAFT ALTERNATORS, C. J. Breitwieser. *AIEE Transactions (Electrical Engineering)*, vol. 64, Nov. 1945, pp. 763–68.

◆

No Discussion

Development of an Air-Borne Stabilized Camera Mount

J. H. MILLER
ASSOCIATE MEMBER AIEE

A. J. ALEXANDER
NONMEMBER AIEE

DURING flight tests at Goodyear Aircraft Corporation it was necessary to ascertain the exact location of the airplane with respect to its plotted course. For this purpose a camera was so mounted in the airplane that it was stabilized with respect to earth co-ordinates in three axes while the airplane was undergoing turns and random maneuvers. Two types of servo (stabilization systems) were considered, a motor-and-gear drive system, and a rate gyro system. The motor-and-gear system proved the more suitable. A torque source drive system is presently being investigated. A prototype has been built, and testing has begun.

The flight-test group required information as to the exact location of an airplane with respect to its plotted course. Terrain photographs made by an air-borne camera supplied this information but it was necessary for the camera to be stabilized vertically and in heading.

Specifications for the stabilized camera mount (see Table I) were set up by the flight-test group to permit camera stability during maneuvers of the airplane used in the tests. However, these specifications are deemed applicable to most present-day aircraft.

A laboratory mock-up simulating one axis of a stabilizing system employing rate gyros was built and tested. Although this system met the error magnitude specifications, it was not completed because adequate rate gyros were not readily available.

Description of Motor and Gear Drive System

GENERAL

Camera stabilization was obtained with a conventional motor-driven position servo. The block diagram of one axis is shown in Fig. 1. The signals for stabilizing the table in the pitch and roll axes were derived from a vertical gyro mounted on the camera table. Deviations between the actual table position and the horizontal were detected by the gyro, amplified, and fed through phase-lead d-c compensating networks to the motor. The compensating networks re-

---

Paper 54-310, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted July 16, 1953; made available for printing April 26, 1954.

J. H. MILLER and A. J. ALEXANDER are with the Goodyear Aircraft Corporation, Akron, Ohio.

Table I. Performance Specifications

	Roll, Degrees	Pitch, Degrees	Azimuth, Degrees
Limits of travel	±30	±20	±25
Max. velocity	90 per second	70 per second	70 per second
Max. acceleration	500 per second^2	150 per second^2	250 per second^2
Accuracy	0.4	0.4	0.5

quired both demodulator and modulator circuits; system stability was not affected by changes in the 400-cycle supply frequency.

The true input to the system is the attitude of the airframe in space. If the airframe changes its attitude in space a given amount on a given axis, the servo must move an equal and opposite amount on the same axis to keep the table angularly stationary in space.

The effects of vertical gyro drift were compensated for by a conventional gyro erection system. Because signals from the erection pick-offs are in error during lateral accelerations, mercury switches were installed on the table to cancel erection when the lateral acceleration exceeded a certain amount.

The azimuth stabilization system was identical to that of the pitch and roll axes except that the error signal was not obtained from the vertical gyro. In azimuth the error signal was obtained from the synchro addition of the airplane's gyrosyn compass position, the position of a course-selector dial on the pilot's control box, and the azimuth position of the camera relative to the airframe.

With this 3-axis system, stabilization of better than 0.15 degree with respect to the gyro was achieved when the camera was subjected to angular velocities of 90 degrees per second. Since the gyro's vertical accuracy was 0.25 degree, the maximum error in space stabilization was less than the specified value of 0.4 degree.

MECHANICAL

The mechanical design of the stabilized camera mount depended upon location of the gyro and the type of servoactuators used. An artist's conception of the camera mount is shown in Fig. 2. All three servos employed spur gearing, torque limit clutches, limit switches, spring stops, and identical motors.

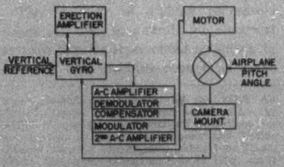


Fig. 1. Block diagram of motor and gear drive servo system

Aluminum was used throughout the structure for a lightweight, yet rigid, design. A casting was used for the stabilized platform; a fabricated box section was used for the gimbal; and a fabricated channel section was used for the outer frame. The weight of the camera mount system, including amplifiers, inverter, and a dynamotor, but excluding the camera, weighed approximately 140 pounds. The camera (an Air Force *K-24*) weighed approximately 26 pounds. The servoamplifiers and a plate-supply regulator were packaged within the outer frame. The erection amplifier and the dynamotor were mounted under this frame making the unit self-contained except for a remote control box and the 28-volt d-c, and 115-volt 400-cycle 3-phase power sources.

Aluminum blocks (heat sinks) were used to prevent overheating in the amplifiers. The blocks surrounded the power tubes and were in direct contact with the outer frame of the camera mount. Blocks and the camera-mount structure were aluminized black to absorb and radiate a maximum amount of heat.

Proposed Torque Source Drive System

A torque source drive system utilizing a magnetic-clutch servo is presently being investigated. An ideal torque source system would employ a pure torque source controller, have no friction, no mass unbalance, and would have all the inertia of the actuating system unity geared to the platform. A pure torque source controller is defined here as one having a torque output independent of the velocity of the load. It can be shown mathematically (see Appendix I) that the position of the stabilized platform in this idealized case is independent of the motion of the airplane; that is, the error in stabilization from the airplane's motion is zero.

With velocity and torque-velocity source controllers the error is a function of the airplane angular velocity since the error must be developed before motion of platform, relative to the airplane, can be produced. The torque source controller, however, permits full utilization of the natural tendency of the platform to stabilize itself in space because of its own inertia, and the stability of the system is not affected by flexure in the actuation and follow-up system (see Appendix II).[1]

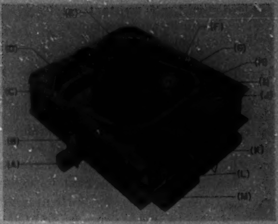


Fig. 2. Artist's conception of a stabilized camera mount: A—Pitch servomotor; B—gear box; C—bubble levels; D—stabilized platform; E—camera; F—vertical gyro; G—gimbal; H—outer frame; I—gear box; J—roll servomotor; K—dynamotor; L—servoamplifiers; M—erection amplifier

Although the idealized state cannot be achieved because the friction cannot be reduced to zero, the gearing ratio between the magnetic clutch and the stabilized platform can be close to unity and the friction can be considerably reduced from that of an electric motor drive system. It is felt that both stabilization error and control power can be reduced with a torque source system. Because of power gain in the magnetic clutch, the electronic amplifier requirements can be greatly reduced, resulting in a higher over-all efficiency and a much smaller amplifier package. The magnetic clutch can be controlled by a single miniature tube, whereas the electric motor system utilizes two tubes comparable to a *6L6*.

A laboratory mock-up of this system is being tested and a hydraulic servo is being designed to obtain a torque source controller by differential pressure feedback.

Appendix I

Symbol Conventions

θ_R = angle from arbitrary reference to horizontal
θ_0 = angle from arbitrary reference to table position
ψ = angle from arbitrary reference to airplane heading axis
λ = angle from airplane heading axis to table position
T = torque
J = inertia
F = damping
K = gain

B = lead-circuit time constant
a = lead-circuit attenuation factor
s = Laplace transform variable

Torque Source System

A system with a controller having the characteristics of a pure torque source may be represented by the following equations

$$T=K\left(\frac{1+saB}{1+sB}\right)\left(\theta_R-\theta_0\right)$$

$$T=s^2J\theta_0$$

$$\theta_0=\lambda+\psi$$

It is assumed that there is no mass unbalance and that all the inertia of the actuating system is unity geared to the platform. Solving the foregoing equations simultaneously yields

$$\theta_0=\left[\frac{1+saB}{1+saB+\frac{s^2J}{K}+\frac{s^3JB}{K}}\right]\theta_R$$

Thus, for the ideal torque source system the table position in space is not a function of motion of the airplane.

Velocity Source System

A system with a controller having the characteristics of a pure velocity source may be represented by the following equations

$$\frac{K}{s}(\theta_R-\theta_0)=\lambda$$

$$\theta_0=\lambda+\psi$$

A pure velocity source is independent of load inertia and damping. Combining these equations and solving for θ_0 yields

$$\theta_0(s)=\left[\frac{1}{1+\frac{s}{K}}\right]\theta_R+\left[\frac{\frac{s}{K}}{1+\frac{s}{K}}\right]\psi$$

Thus, for the ideal velocity source system the table position in space is a function of the airplane's rate of change of heading.

Torque-Velocity Source System

A system with a controller having the characteristics of a torque-velocity source may be represented by the following equations

$$T=K(\theta_R-\theta_0)$$

$$T=s^2J\theta_0+sF\lambda$$

$$\theta_0=\lambda+\psi$$

It is assumed here that there is no mass unbalance and that all the inertia of the actuating system is unity geared to the platform. Combining these equations and solving for θ_0 yields

$$\theta_0(s)=\left[\frac{1}{1+\frac{sF}{K}+\frac{s^2J}{K}}\right]\theta_R+\left[\frac{\frac{sF}{K}}{1+\frac{sF}{K}+\frac{s^2J}{K}}\right]\psi$$

Here also the table position in space is a function of the airplane's rate of change of heading.

Appendix II. Torque Source Stability

Symbol Conventions

θ_R = angle from arbitrary reference to horizontal
θ_0 = angle from arbitrary reference to table position
ψ = angle from arbitrary reference to airplane heading axis
λ = angle from airplane heading axis to table position
T = torque
J = inertia
F = damping
K = gain
B = lead-circuit time constant
a = lead-circuit attenuation factor
s = Laplace transform variable
λ_1 = angle from airplane heading to output of actuating source
λ_2 = angle from airplane heading to table position
K_x = spring constant of actuating system due to flexure

Torque Source Stability

A system with flexure in the actuating system and with a controller having the characteristics of a pure torque source may be represented by the following equations

$$T=K\left(\frac{1+saB}{1+sB}\right)(\theta_R-\theta_0)$$

$$T=K_x(\lambda_1-\lambda_2)$$

$$\theta_0=\lambda_2+\psi$$

$$0=s^2J\theta_0+K_x(\lambda_2-\lambda_1)$$

Although a practical system with flexure would be in general a distributed constant system, a lumped constant system is assumed here for simplicity. Also, it is assumed that there is no mass unbalance and that all the inertia of the actuating system is concentrated at the table. Solving the above equations simultaneously yields

$$\theta_0=\left[\frac{1+saB}{1+saB+\frac{s^2J}{K}+\frac{s^3JB}{K}}\right]\theta_R$$

The result obtained is identical to that obtained in Appendix I. Therefore, the stability and performance of the torque source system is independent of the flexure in the actuating system.

Reference

1. Engineering Bulletin No. 6. Standard Controls, Inc., Seattle, Wash., Dec. 1952.

◆

No Discussion

Transient Characteristics of Aircraft A-C Generators

V. C. HOLLOWAY
MEMBER AIEE

DURING the past 5 years transient analysis of aircraft a-c generators has become increasingly important. The vast number of excellent technical papers[1–3] published on this subject would indicate that no problem existed, and a literature survey would yield a solution to any problem no matter what the degree of complexity. However, problems do exist in transforming from conventional power frequencies to 400 cycles per second (cps) aircraft equipments and in ascertaining the idealizations which other authors used in the formulation of the various papers.

How can an engineer evaluate the transient behavior of an a-c generator? Various authors[4] have stated in essence, "Determine the reactances, and the complete behavior of the machine under any condition is known." It is the purpose of this paper to enumerate the various methods of obtaining the synchronous reactances and resistances and to suggest practical test methods to be used for the types of machines used at present.

For 60-cps a-c generators, reliable constants are available for a reasonable system evaluation; however, for 400-cps aircraft equipments, this is generally not true. This work has been conducted in order to afford aircraft electrical system design groups sufficient information prior to the delivery of a-c generators for various system predictions. These include maximum interrupting currents for circuit breakers, maximum phase voltages resulting from transient switching and current or voltage amplitudes, and durations for prescribed unbalanced conditions.

Idealized 3-Phase Generator

The majority of 3-phase generators which are used in aircraft a-c electrical systems may be represented as shown in Fig. 1. This machine consists of three stationary winding symmetrically displaced and two additional windings rotating with respect to the aforementioned stationary windings. An equivalent circuit diagram for a salient-pole synchronous machine[5] with amortisseur winding is shown in Fig. 2. The problem evolved to find the various constants shown in Fig. 2 and to obtain the steady-state and transient solutions therefrom.

Units

In this paper, only per-unit values are used. In the analytical determination of certain reactances, it is usually advantageous to use per-unit values. The ohmic value of reactances and resistances obtained by various test methods can be converted to per-unit values by multiplying by rated phase current and dividing by rated phase voltage.

Description of Test Methods

SYNCHRONOUS REACTANCE

X_d—The synchronous reactance X_d can be obtained by method 1, paragraph 1.822 of the AIEE Test Code for Synchronous Machines.[6]

QUADRATURE REACTANCE

X_q—The quadrature reactance X_q on power generators is usually obtained by the so-called slip test. However, on relatively small rating aircraft machines, it is quite difficult to obtain repeatable results using this method. Large errors are introduced by the effects of currents induced by amortisseur circuits unless the slip can be made very small. Method 1, paragraph 1.832 of the AIEE Test Code for Synchronous Machines can be used to obtain the quadrature reactance X_q. Refer to Fig. 3, $X_q = E/I$, where X = per-unit armature phase voltage and I = per-unit phase current at the maximum stable negative excitation.

ARMATURE LEAKAGE REACTANCE

X_1—The armature leakage reactance is defined as the difference between the synchronous reactance X_d and the reactance of armature reaction X_{ad}. Many papers have been written pointing out the differences between the Potier reactance, which can easily be tested for and the armature leakage reactance. The Potier reactance (X_p) varies with magnetic loading and is obtained by method (a), paragraph 1.520, AIEE Test Code for Synchronous Machines.

Since the Potier reactance more nearly approaches the leakage reactance at high values of field currents, the method for determination of the Potier reactance should be extended to 3 or 4 times per-unit field current; 1 per unit field current is arbitrarily defined as the field current to produce rated voltage on the air-gap line. Fig. 4 shows the variance of the Potier reactance as a function of terminal voltage.

DIRECT-AXIS ARMATURE REACTION REACTANCE

X_{ad}—Two quantities for the direct axis have been evaluated, namely, X_d and X_1. Therefore

$$X_{ad} = X_d - X_1 \tag{1}$$

QUADRATURE-AXIS ARMATURE REACTION REACTANCE

X_{aq}—Similarly, two quantities have been determined for the quadrature-axis, namely, X_q and X_1. Therefore

$$X_{aq} = X_q - X_1 \tag{2}$$

FIELD LEAKAGE REACTANCE

X_f—To obtain the field leakage reactance X_f, design calculations are made in accordance with formulas developed by either Alger[7] or Kilgore.[8]

AMORTISSEUR DIRECT-AXIS AND QUADRATURE-AXIS REACTANCES

X_{kd} and X_{kq}—The amortisseur direct-axis and quadrature-axis reactances are obtained also from design calculations.[7,8] Design calculations imply performance predictions from winding data, materials and machine configurations.

The various reactances shown in Fig. 2

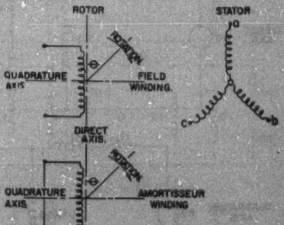


Fig. 1. Idealized 3-phase generator

Paper 54-305, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted June 30, 1953; made available for printing April 23, 1954.

V. C. HOLLOWAY is with the Naval Research Laboratory, Washington, D. C.

Nomenclature

X_d = direct-axis synchronous reactance
X_d' = direct-axis transient reactance
X_d'' = direct-axis subtransient reactance
X_{ad} = direct-axis armature reaction reactance
X_q = quadrature-axis synchronous reactance
X_q' = quadrature-axis transient reactance
X_q'' = quadrature-axis subtransient reactance
X_{aq} = quadrature-axis armature reaction reactance
X_2 = negative-sequence reactance
X_0 = zero-sequence reactance
X_1 = armature leakage reactance
X_p = Potier reactance
r_a = effective armature resistance
r_{am} = amortisseur resistance
x_f = field leakage reactance
r_f = field resistance
x_{kd} = direct-axis amortisseur reactance
r_{kd} = direct-axis amortisseur resistance
x_{kq} = quadrature-axis amortisseur reactance
r_{kq} = quadrature-axis amortisseur resistance
T_d' = direct-axis transient short-circuit time constant
T_d'' = direct-axis subtransient short-circuit time constant
T_q' = quadrature-axis transient short-circuit time constant
T_q'' = quadrature-axis subtransient short-circuit time constant
T_{do}' = direct-axis transient open-circuit time constant
T_a = direct-axis armature time constant

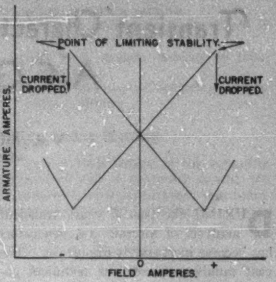


Fig. 3. Maximum lagging current test

have now been evaluated. To obtain various transient reactances, the field and amortisseur quantities are referred to the stator. Neglecting circuit resistances, from Fig. 5(A), the direct-axis transient reactance X_d' is

$$X_d' = X_1 + \cfrac{1}{\cfrac{1}{X_{ad}} + \cfrac{1}{X_f}} \tag{3}$$

Similarly, from Fig. 5(B), the direct-axis subtransient reactance X_d'' is

$$X_d'' = X_1 + \cfrac{1}{\cfrac{1}{X_{ad}} + \cfrac{1}{X_f} + \cfrac{1}{X_{kd}}} \tag{4}$$

From Fig. 5(C), the quadrature-axis subtransient reactance X_q'' is

$$X_q'' = X_1 + \cfrac{1}{\cfrac{1}{X_{aq}} + \cfrac{1}{X_{kq}}} \tag{5}$$

After the amortisseur currents die out, the quadrature-axis transient and quadrature-axis synchronous reactances are equal.

$$X_q' = X_q = X_1 + X_{aq} \tag{6}$$

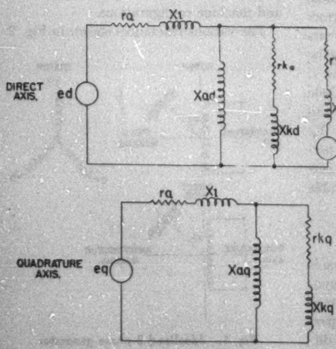


Fig. 2 (left). Equivalent circuit for a salient-pole a-c generator

Negative Sequence Reactance

X_2—The negative sequence reactance X_2 can be obtained by method 1, paragraph 1.883, AIEE Test Code for Synchronous Machines.

Zero Sequence Reactance

X_0,—The zero sequence X_0 can be obtained by method 2, paragraph 1.894, AIEE Test Code for Synchronous Machines.

Resistance Test Methods

Effective Armature Resistance

r_a—The effective armature resistance r_a is the power loss associated with the armature current divided by the square of the current. The a-c generator is run at synchronous speed and the terminals carefully short-circuited through ammeters. The torque input is measured by means of a dynamometer. The excitation is slowly increased from zero. A curve similar to Fig. 6 can be plotted from the results. Referring to the direct-axis synchronous impedance curve, the relationship between armature current and field excitation can be noted. The core loss, which is included in the I^2r_a loss, is usually neglected due to low saturation. The effective armature resistance r_a, which is independent of the magnitude of armature current is usually determined for rated line current.

Field Resistance

r_f—The field resistance r_f can be evaluated from design data or measured directly if the test machine is available. The measured or calculated value of field resistance r_f must then be referred to the stator circuit (Fig. 2).[9]

Amortisseur Resistance

r_{am}—The amortisseur resistance r_{am} can be obtained from design calculations,[7] assuming the amortisseur circuit to be a squirrel-cage winding. The individual amortisseur bars cannot be separated discretely into two groups of bars; one associated with the direct-axis; the other, with the quadrature axis. For a more complete analysis of amortisseur circuits, references 5 and 10 should be consulted.

Time Constants

Any current which decreases exponentially with respect to time from an ini-

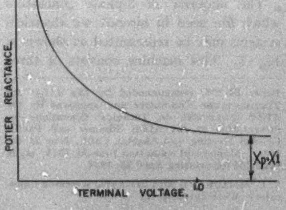


Fig. 4 (right). Variance of Potier reactance versus terminal voltage

Table I. Comparison of Calculated and Test Results for Various Aircraft Generators at 400 Cps

Quantity	Generator A, 15 Kva, 6,000 Rpm		Generator B, 15 Kva, 6,000 Rpm		Generator C, 30 Kva, 6,000 Rpm		Generator D, 30 Kva, 4,000 Rpm		Generator E, 60 Kva, 6,000 Rpm	
	Calc.	Test	Calc.	Test	Calc.	Test	Calc.	Test	Calc.	Test
X_{ad}	1.40	1.30	1.14	*	2.14	2.15	1.21	1.26	3.02	‡
X_l	0.14	0.15	0.095		0.205	0.185	0.147	0.158	0.153	
X_d	1.54	1.45	1.24		2.35	2.34	1.36	1.42	3.17	
X_d'	0.28	0.27	0.215		0.42	0.45	0.302	0.263	0.486	
X_d''	0.17	0.23	0.12		0.254	0.28	0.226	0.170	0.23	
††X_d''†		0.21				0.26		0.20		
X_{aq}	0.76	0.87	0.60		1.14	1.31	0.638	0.61	1.58	
X_q	0.90	1.02	0.70		1.35	1.47	0.785	0.77	1.73	
X_q''	0.18		0.127		0.276		0.29		0.253	
††X_q''†		0.25				0.32		0.23		
X_2	0.18	0.21	0.124		0.265	0.25	0.258	0.215	0.24	
X_0	0.06	0.08	0.003		0.125	0.10	0.070	0.056	0.104	
X_f	0.158		0.134		0.235		0.178		0.375	
X_{kd}	0.036		0.031		0.063		0.162		0.096	
X_{kq}	0.044		0.034		0.073		0.184		0.11	
r_a	0.056	0.06	0.032		0.044	0.044	0.058	0.06	0.029	
r_f	0.028		0.014		0.023		0.036		0.017	
r_{am}	0.034		0.022		0.042		0.057		0.063	
T_d''	0.0013	0.001	0.0015		0.0016	0.0012	0.0017	0.0010	0.0013	
T_d'	0.004	0.0033	0.0075		0.0073	0.0083	0.0034	0.0046	0.0034	
T_a	0.0013	0.0019	0.0015		0.0024	0.0022	0.0018	0.0018	0.0033	
T_{do}'	0.064		0.011		0.116		0.096		0.228	

* Generator not available for tests.

† Static test in accordance with reference 13.

‡ Test evaluation not completed.

tial value of I_0 at time $t=0$ to a fixed value of zero can be expressed as a function of time as

$$i=I_0e^{-t/T} \tag{7}$$

In a synchronous machine, the resistance and reactances are as shown in Fig. 2. Various time constants may now be evaluated.[11] Referring to Fig. 7(A), the direct-axis short-circuit time constant T_d' is, neglecting resistances in all parallel branches

$$T_d'=\frac{X_f+\dfrac{1}{\dfrac{1}{X_{ad}}+\dfrac{1}{X_f}}}{2\pi f\,r_f} \tag{8}$$

From Fig. 7(B), the direct-axis short-circuit subtransient time constant T_d'' is as follows

$$T_d''=\frac{X_{kd}+\dfrac{1}{\dfrac{1}{X_f}+\dfrac{1}{X_{ad}}+\dfrac{1}{X_l}}}{2\pi f\,r_{kd}} \tag{9}$$

From Fig. 7(C), the quadrature-axis short-circuit subtransient time constant is

$$T_q''=\frac{X_{kq}+\dfrac{1}{\dfrac{1}{X_{aq}}+\dfrac{1}{X_l}}}{2\pi f\,r_{kq}} \tag{10}$$

Direct-axis open-circuit subtransient time constant T_{do}'. Referring to Fig. 7(D)

$$T_{do}'=\frac{X_f+X_{ad}}{2\pi f\,r_f} \tag{11}$$

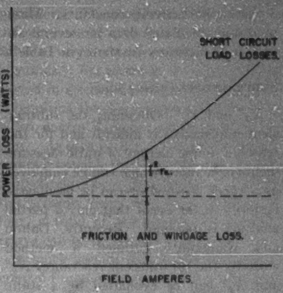


Fig. 6. Power curve for short-circuit test

ARMATURE SHORT-CIRCUIT TIME CONSTANT

T_a—The armature short-circuit time constant T_a at the machine terminals is dependent upon the effective armature resistance r_a and the inductance of the armature circuit to direct current. From Fig. 2, the rotor circuits are inductively coupled to the stator and depend upon angular displacement for discrete values. Usually the arithmetic mean of X_d'' and X_q'' is used (i.e., X_2). Therefore

$$T_a=\frac{\dfrac{X_d''+X_q''}{2}}{2\pi f\,r_a}=\frac{X_2}{2\pi f\,r_a} \tag{12}$$

3-Phase Short-Circuit Constants

The analysis of a 3-phase short-circuit test enables X_d'', X_d', T_d'', T_d', and T_a to be evaluated. Wright[13] described methods of determining these constants by test. A manual used by the writer outlines in considerable detail the test methods to be used for evaluation of the

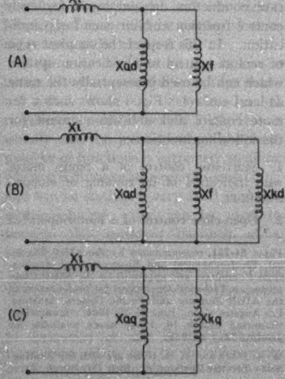


Fig. 5 (left). Equivalent circuits for determining transient reactances

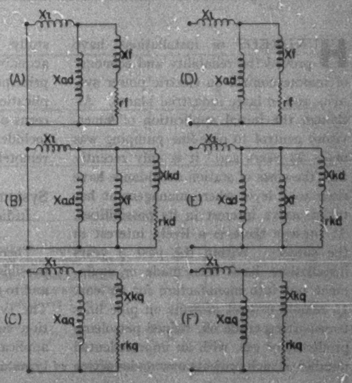


Fig. 7 (right). Equivalent circuits for determining open-circuit and short-circuit time constants

3-phase short-circuit constants. Measured and calculated data for several aircraft a-c generators are shown in Table I.

SUBTRANSIENT IMPEDANCES

X_d'' and X_q''—Obtaining the subtransient impedances is difficult and for this reason a simple method of static measurements is included which can be employed as a check method or can be used satisfactorily in the event that only a partial analysis is required. Recently, Dalton and Cameron[13] outlined a method whereby the subtransient reactances X_d'' and X_q'' can be evaluated by a static test. Comparison of results are shown in Table I.

Design Calculations

System design often precedes component design in aircraft electrical planning. Considerable system evaluation may be accomplished if realistic generator constants are made available to system design groups prior to the actual delivery of an approved generator. Generator manufacturers have made available to the writer sufficient design information so that calculations could be made as outlined by previous writers. From these data certain specification requirements such as voltage unbalance and efficiency have been accurately determined and verified by later tests. Table I shows correlation between calculated and test results for generator constants. This table shows agreement within 10 per cent or less between measured and calculated values.

Conclusions

Outlined in the foregoing are methods for the evaluation of the constants of an aircraft a-c generator. For the definitive impedances, those measured from the stator terminals (X_d, X_{ad}, X_1, r, etc.) standard test procedures are outlined. For the nondefinitive impedances, those impedances which cannot be measured directly from the stator terminals (X_f, X_{kd}, X_{kq}, r_{kd}, etc.) methods have been outlined for evaluation and then transformation to equivalent stator values. After obtaining a direct-axis and quadrature-axis, equivalent circuit, open and short-circuit time constants have been determined. In addition, design calculations are compared with actual test results for reactances and time constants.

References

1. TRANSIENT ANALYSIS OF A-C MACHINERY, **Yu H. Ku.** *AIEE Transactions*, vol. 48, July 1929, pp. 707–15.

2. TWO-REACTION THEORY OF SYNCHRONOUS MACHINES—I, **R. H. Park.** *AIEE Transactions*, vol. 48, July 1929, pp. 717–30.

3. TWO-REACTION THEORY OF SYNCHRONOUS MACHINES—II, **R. H. Park.** *AIEE Transactions*, vol. 52, June 1933, pp. 353–54.

4. SYNCHRONOUS MACHINE REACTANCES, **L. P. Shildneck.** *General Electric Review*, Schenectady, N. Y., Nov. 1932.

5. EQUIVALENT CIRCUIT OF THE SALIENT-POLE SYNCHRONOUS MACHINE, **Gabriel Kron.** *General Electric Review*, Schenectady, N. Y., Dec. 1941.

6. TEST CODE FOR SYNCHRONOUS MACHINES. *AIEE Standard No. 503*, June 1945.

7. THE NATURE OF POLYPHASE INDUCTION MACHINES (book), **P. L. Alger.** John Wiley and Sons, Inc., New York, N. Y., 1951.

8. CALCULATION OF SYNCHRONOUS MACHINE CONSTANTS, **L. A. Kilgore.** *AIEE Transactions*, vol. 50, Dec. 1931, pp. 1201–14.

9. DESIGN OF ELECTRICAL APPARATUS (book), **J. H. Kuhlman.** John Wiley and Sons, Inc., New York, N. Y., third edition, 1950.

10. STARTING PERFORMANCE OF SALIENT-POLE SYNCHRONOUS MOTORS, **T. M. Linville.** *AIEE Transactions*, vol. 49, Apr. 1930, pp. 531–47.

11. CIRCUIT ANALYSIS OF A-C POWER SYSTEMS (book), **Edith Clarke.** John Wiley and Sons, Inc., New York, N. Y., vol. II, 1950.

12. DETERMINATION OF MACHINE CONSTANTS BY TEST, **S. H. Wright.** *AIEE Transactions* (*Electrical Engineering*), vol. 50, no. 4, Dec. 1931, pp. 1331–51.

13. SIMPLIFIED MEASUREMENT OF SUBTRANSIENT AND NEGATIVE SEQUENCE REACTANCES IN SYNCHRONOUS MACHINES, **F. K. Dalton, A. W. W. Cameron.** *AIEE Transactions*, vol. 71, pt. III, Feb. 1952, pp. 753–57.

◆

No Discussion

Remote Operation of Pipe-Line Pumping Stations

W. A. DERR
MEMBER AIEE

M. A. HYDE
MEMBER AIEE

HUNDREDS of installations have proved the reliability and economy of remote control on electric power systems and in large industrial plants. Although the initial application of supervisory control to pipe-line pumping was made 23 years ago,[1] it is only recently that the costs of station attendance have risen to a level where management has taken active interest in its possibilities. At present there is a lively interest in the subject. Within the past 3 years installations have been made or equipment put into manufacture for at least 15 remote control projects on pipe lines transporting crude oil, refined petroleum products, or gas, with an unprecedented number of such projects now under active study. This paper is written to better acquaint pipe-line engineers with the principles of remote control and its application to pipe-line service. It concerns only remote-control systems which include indication of the condition of the remotely operated equipment.

Systems for Remote Control and Indication

There are many ways in which it is possible to control apparatus remotely and to provide indications of its positions. The type of remote-control and indication equipment chosen for a particular application is usually determined by the types and costs of channels which can be made available and by the degree of reliability deemed necessary. Some of the basic methods of remote control and indication which can be employed will first be briefly discussed.

DIRECT CONTROL AND INDICATION

For local control and indication separate conductors are employed for each control function and for each lamp indication. In this respect the simplest type of remote-control and indication system which can be used is essentially the same as local control. Fig. 1 shows such a remote control and indication circuit for the following functions:

1. Start-stop control of a pump motor with indication of its running or stopped condition.

2. Open-close control of a motor-operated

Paper 54-221, recommended by the AIEE Chemical, Electrochemical and Electrothermal Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 16, 1954; made available for printing May 4, 1954.

W. A. DERR and M. A. HYDE are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

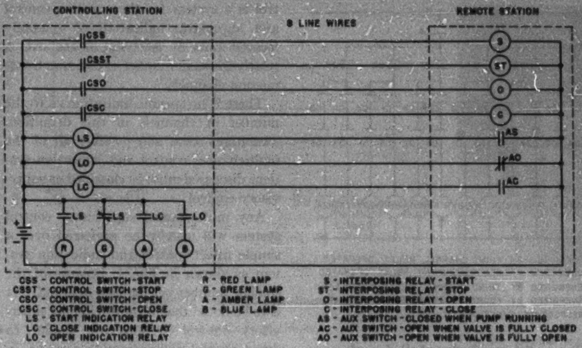

CONTROLLING STATION
8 LINE WIRES
REMOTE STATION
CSS
CSST
CSO
CSC
LS
LO
LC
S
ST
O
C
AS
AO
AC
R
G
A
B
CSS - CONTROL SWITCH - START
CSST - CONTROL SWITCH - STOP
CSO - CONTROL SWITCH - OPEN
CSC - CONTROL SWITCH - CLOSE
LS - START INDICATION RELAY
LC - CLOSE INDICATION RELAY
LO - OPEN INDICATION RELAY
R - RED LAMP
G - GREEN LAMP
A - AMBER LAMP
B - BLUE LAMP
S - INTERPOSING RELAY - START
ST - INTERPOSING RELAY - STOP
O - INTERPOSING RELAY - OPEN
C - INTERPOSING RELAY - CLOSE
AS - AUX SWITCH - CLOSED WHEN PUMP RUNNING
AC - AUX SWITCH - OPEN WHEN VALVE IS FULLY CLOSED
AO - AUX SWITCH - OPEN WHEN VALVE IS FULLY OPEN

Fig. 1. Simple direct control and indication for one pump and one valve over multiple channels

valve with indication of its open or closed condition.

As shown in Fig. 1, a total of eight conductors must be provided between the controlling and remote stations for the two complete functions listed. The number of conductors required with such a scheme becomes extremely large when a considerable number of functions must be performed, because individual conductors are used for each function.

It is apparent that other types of individual channel means could be substituted for the line wires shown in Fig. 1. For example, seven individual telegraphic-type channels[7] could be provided over a single pair of wires to perform the functions shown. The seven telegraphic-type channels could be obtained by seven sets of signal or tone-generating and tone-receiving equipment in the audio range. Likewise, individual audio tones could be used to modulate telephone or power-line carrier channels to perform these functions, or individual audio or higher frequency tones could be employed on a microwave radio channel.

This system of direct control and indication is very economical except for channel costs. However, since the channel requirements are extremely high, it can usually only be justified when the total number of functions is extremely small or where the distance between the controlling and remote stations is very short so that a multiconductor cable can be used to provide the necessary channels at a reasonable cost.

The simple scheme shown in Fig. 1 has two serious disadvantages from the standpoint of reliability:

1. Operation of any of the individual interposing relays due to the momentary presence of a disturbing voltage on the interconnecting channel results in an unwanted operation of a controlled device.
2. It is impossible to determine whether or not the individual channels are in working condition until they must be used.

The fact that the individual channels cannot be continuously supervised means that a channel may not be available when it is needed to perform a function.

Selective Control and Indication Using Multiple Channels

In Fig. 1, four conductors in addition to a common conductor are used to operate a total of four interposing control relays. The four control wires are used one at a time and only during the instant that a control contact is closed to perform an operation. These otherwise inactive control wires can be employed in a number of different ways to obtain additional control functions. For example, they can be employed in a scheme based on the principle of requiring the simultaneous energization of two control wires for each control function. As seen in Fig. 2 this arrangement makes it possible to operate six interposing control relays over the four wires. With this scheme, using five wires instead of four, a total of ten interposing control relays can be individually operated.

A much greater gain in the number of control functions performed over a given number of channels is obtained by the use of a system based on the sequential use of all of the channels provided. For example, if four control wires in addition to a common wire are used, circuitry can be provided at the controlling station so that each of the four wires is momentarily energized in a definite and different sequence for each control function. Likewise, circuitry can be provided at the remote station to recognize the order in which the wires are energized and finally to energize a particular interposing control relay dependent on the order in which the wires were energized. If such a sequential scheme is employed, using four channel-operated relays, 24 different interposing control relays can be individually operated as shown in Fig. 3. The scheme is based on one of the *A* relays being energized by the first channel relay operated, one of the *B* relays by the second channel relay operated, etc. If the channel-operated relays are energized in the order 4-3-2-1, sequence relays *A4*,

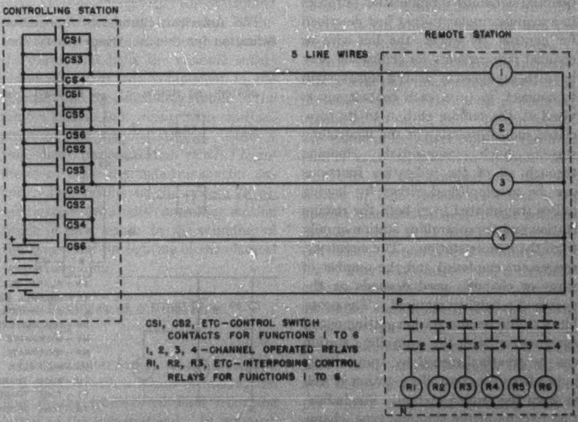

CONTROLLING STATION
REMOTE STATION
5 LINE WIRES
CS1
CS3
CS4
CS1
CS5
CS6
CS2
CS3
CS5
CS2
CS4
CS6
1
2
3
4
P
N
R1
R2
R3
R4
R5
R6
CS1, CS2, ETC — CONTROL SWITCH CONTACTS FOR FUNCTIONS 1 TO 6
1, 2, 3, 4 — CHANNEL OPERATED RELAYS
R1, R2, R3, ETC — INTERPOSING CONTROL RELAYS FOR FUNCTIONS 1 TO 6

Fig. 2. Multiple-channel remote control by simultaneous energization of two channels

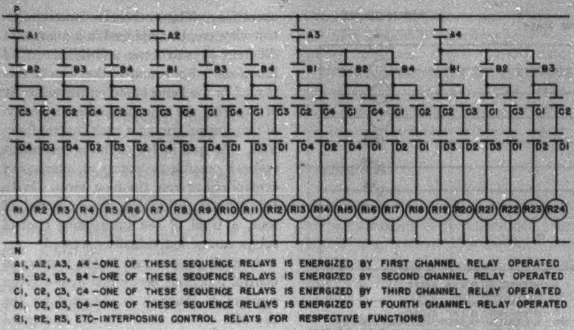

P
A1
A2
A3
A4
B1
B2
B3
B4
C1
C2
C3
C4
D1
D2
D3
D4
R1 R2 R3 R4 R5 R6 R7 R8 R9 R10 R11 R12 R13 R14 R15 R16 R17 R18 R19 R20 R21 R22 R23 R24
N
A1, A2, A3, A4 – ONE OF THESE SEQUENCE RELAYS IS ENERGIZED BY FIRST CHANNEL RELAY OPERATED
B1, B2, B3, B4 – ONE OF THESE SEQUENCE RELAYS IS ENERGIZED BY SECOND CHANNEL RELAY OPERATED
C1, C2, C3, C4 – ONE OF THESE SEQUENCE RELAYS IS ENERGIZED BY THIRD CHANNEL RELAY OPERATED
D1, D2, D3, D4 – ONE OF THESE SEQUENCE RELAYS IS ENERGIZED BY FOURTH CHANNEL RELAY OPERATED
R1, R2, R3, ETC – INTERPOSING CONTROL RELAYS FOR RESPECTIVE FUNCTIONS

Fig. 3. Multiple-channel remote control by sequential energization of all channels

B3, *C2*, and *D1* will be energized. This results in the operation of interposing control relay *R24*.

The indication problem differs somewhat from the control problem. For control it is only necessary to operate a particular control switch. For indication it is necessary to position an indication relay at the controlling station to correspond to the position of each device supervised, and to have these indication relays automatically change position whenever there is a change in the position of a device supervised. There are many different component and circuit arrangements which can be employed to provide indications with fewer channels than the maximum of one per indication shown in Fig. 1. Some of these will be briefly discussed.

It is possible to use simultaneous or sequential signalling for indication purposes in a manner similar to that just described for control. However, the line-wire or channel requirements for indication may be further reduced by using a pulse system to connect, in turn, each contact supervised at the remote station to its associated indication relay at the controlling station. Such sequential stepping through all of the indication functions can be accomplished either by having pulses transmitted from both the remote station and the controlling station or only from the remote station. The circuit arrangement employed and the number of wires or channels used depends on the degree of reliability required. The possibility of the remote station getting out of synchronism with the controlling station can be greatly lessened by transmitting pulses from the controlling station as well as from the remote station. For indication circuits of the stepping type, chains of telephone-type relays or telephone-type selector switches can be employed. Less maintenance at a somewhat higher initial cost is realized with all-relay systems.

Multichannel systems have been designed which use the same channels for control and indication.[3] These effect a further reduction in channel requirements.

In the foregoing discussion of multichannel systems line wires have been considered as the interconnecting channels. It is apparent, however, that other telegraphic-channel means could be substituted for the line wires, as was discussed for the direct control and indication scheme. Any means which provides for the individual energization of a relay at the remote location is considered a channel.

Supervisory Control

The American Standards Association definition for this is: "Supervisory control is a system for the selective control and automatic indication of remotely located units by electrical means, over a relatively small number of common transmission channels."

There is no specific limitation as to the number of channels in this definition. Therefore, some and possibly all of the multichannel control and indication systems discussed must be classified as supervisory control.

Any multichannel supervisory control system will usually be preferred over a simple direct-control and indication system as shown in Fig. 1, because of the smaller number of channels required. It is to be noted that all multichannel supervisory control systems are similar to direct control and indication systems, in that it is impossible to supervise continuously all of the channels required for the operation of the system.

Single-Channel Supervisory Control

Supervisory control systems which require the minimum of a single channel-operated relay at both the controlling station and the remote station have been available for more than 20 years. Such systems must obviously be of the pulse type to transmit all necessary information. With a single channel-operated relay, differentiation for selection can only be obtained by varying the number of pulses, by varying the length of the pulses, or by varying the length of the pauses between pulses.[4–7] Systems of this general type must either provide a self-checking type of coding or a verification or check code to preclude the possibility of incorrect selection. Such systems have been manufactured using all telephone-type relays or a combination

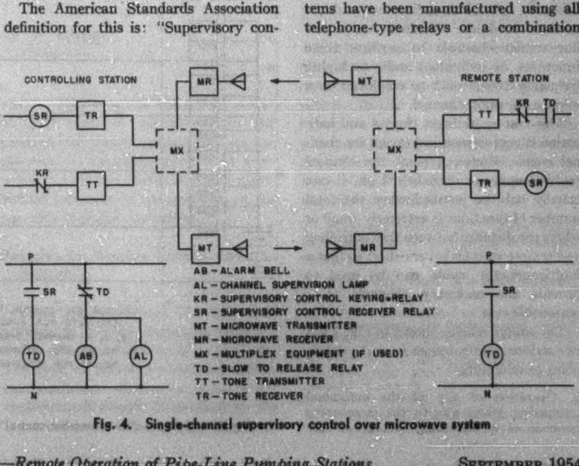

CONTROLLING STATION
REMOTE STATION
MR
MT
MT
MR
SR
TR
KR
TT
MX
MX
TT
KR
TD
TR
SR
P
SR
TD
TD
AB
AL
N
P
SR
TD
N
AB – ALARM BELL
AL – CHANNEL SUPERVISION LAMP
KR – SUPERVISORY CONTROL KEYING RELAY
SR – SUPERVISORY CONTROL RECEIVER RELAY
MT – MICROWAVE TRANSMITTER
MR – MICROWAVE RECEIVER
MX – MULTIPLEX EQUIPMENT (IF USED)
TD – SLOW TO RELEASE RELAY
TT – TONE TRANSMITTER
TR – TONE RECEIVER

Fig. 4. Single-channel supervisory control over microwave system

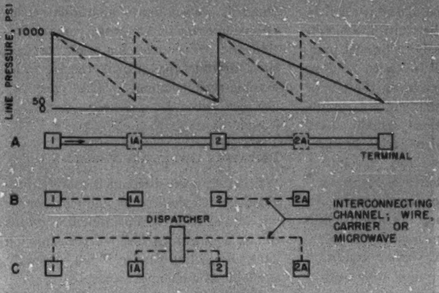


Fig. 5. Diagram of simplified pipe-line system

A—Typical station sequence and pressure gradients
B—Two-station-link supervisory control
C—Centralized system control from dispatching office

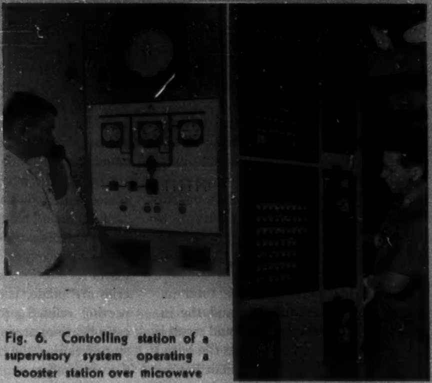


Fig. 6. Controlling station of a supervisory system operating a booster station over microwave

Left: Operating panel, with push buttons, indicating lights and pressure indicators in diagram; flow recorder is above this
Right: Rear of operating panel, showing telephone-type supervisory control relays, and frequency-actuated telemetering receivers for three pressures and flow

of telephone-type relays and selector switches

In these systems it is customary to use similar codes for control and for indication, the same chains of relays or selector switches being employed for the registering of all control and indication pulses. The indications are considered to be just as important as the control functions, so the systems are inherently designed to provide for both control and indication functions. They do not consist of two basically separate systems for control and indication as is the case with many multichannel systems.

Complete channel supervision can be provided with single-channel supervisory control when the system is operated by d-c pulses over a pair of line wires, or when the system is operated over a duplex channel. In supervisory control terminology a simplex channel is one in which transmitted pulses actuate the channel-operated relay at both the transmitting and receiving stations. A duplex channel is one in which transmitted pulses actuate only the channel-operated relay at the receiving station.

For duplex-channel operation the tones or other channel means are transmitted with the equipment at rest; pulses constitute interruption of the channel means. The system is usually arranged so that a failure either in the channel means from the controlling station to the remote station, or in the channel means from the remote station to the controlling station, results in audible and visual alarms at the controlling station.

The fact that single-channel systems can provide complete channel supervision is particularly advantageous where audio or higher frequency tones are used for modulation on a microwave channel, because such channel supervision includes the components of the tone channel as well as the microwave channel itself.

The necessary channel components for operation of single-channel supervisory control over a microwave system are shown in Fig. 4. The tone transmitters at both the controlling station and the remote station are normally keyed by break contacts of the keying relay *KR*. Pulses constitute interruption of the tone by energization of the keying relay *KR*. The arrangement shown provides for the complete supervision of the duplex supervisory control channel including the tone transmitters and receivers and multiplexing equipment as well as the microwave transmitters and receivers. If any component in the channel from the remote station to the controlling station fails, relay *SR* is de-energized at the controlling station. After a time delay relay *TD* releases at the controlling station to sound the alarm and light the channel supervision lamp. If any component in the channel from the controlling station to the remote station fails, relay *SR* is de-energized at the remote station. After a time delay remote-station relay *TD* releases to open the tone-transmitter circuit. This results in the release of receiver relay *SR* at the controlling station which in turn results in the sounding of the alarm and the lighting of the channel supervision lamp.

Basic Features of a Pipe-Line System

In its simplest form a pipe line consists of a conduit with a means of causing the flow of the transported fluid from the initial to the terminal end. In practice, flow is caused by pressure developed by a pump at the initial end. In a crude oil or product line, for a given liquid and pipe size, each flow rate requires the application of a certain pressure per mile of pipe, and the total pressure that must be developed by the pump is the pressure drop per mile multiplied by the length of pipe in miles, plus the positive or negative static pressure corresponding to the difference in elevation of the two ends of the line. Thus, the length of line that can be served by one pumping station is determined by the desired flow rate, the difference in elevation of the ends, and the limiting working pressure for the pipe.

To accomplish long-distance transportation several pipe-line sections are installed in series, each section with a pumping station at its initial end. In the past it was common to deliver from the terminal end of each section into a tank from which the oil was pumped into the next section of line. Modern practice is to operate the sections end-to-end without tankage. Fig. 5(A) shows diagrammatically such a system in level terrain consisting of a series of stations and a terminal. Consider first that the system comprises an initial station no. 1 and one downstream station no. 2. In Fig. 5(A) the pressure conditions are shown in solid line. Each station develops the allowable discharge pressure, say 1,000 pounds per square inch (psi), working with a suction pressure of, say, 50 psi. At design flow the line pressure decreases along a straight-line gradient to the value of 50 psi at the next downstream station.

The throughput of a line can be increased by employing a greater total impelling pressure. However, since the discharge pressure at stations no. 1 and no. 2 is limited to 1,000 psi by allowable pipe stress, the practical solution to obtain additional throughput is to insert pumping stations at intermediate points between the original stations. Such booster stations may develop any pressure up to the allowable line pressure. In Fig. 5, *1A* and *2A* indicate booster stations at midpoints developing the same pressure as the original stations. The total impelling pressure is doubled, and the increased flow is the same as would result if this pressure were all applied at the initial station, the new pressure gradient having a steeper slope (greater psi drop per mile) as shown by the broken line in Fig. 5(A).

Operation in 2-Station Links

The most frequent type of remote operation is the 2-station link shown in Fig. 5(B), where booster stations *1A* and *2A* are remotely operated, each from a neighboring attended station. Usually the controlling station is upstream, because from changes in his local operation the attendant there can readily anticipate the conditions which will follow at the next station downstream. This arrangement is particularly applicable where electric booster stations are inserted between existing stations for increased line capacity. In a series of such stations over level terrain, a throughput increase of about 50 per cent is attained by doubling the number of stations without increasing the number of station operators. This saving in attendant personnel is often a decisive factor in justifying such an expansion in the face of increased investment, power cost, and system maintenance.

Centralized Remote Control

This arrangement is one in which the control for a complete series of stations is carried out from a single location, as shown in Fig. 5(C). From the equipment standpoint the controlling location could be the initial station or the terminal, but if the system is to be centrally controlled it is logical to place this responsibility directly under the dispatching personnel. This location is normally at the administrative headquarters, which is likely to be near the center of the line. Although at present the number of such systems is very limited, it appears that future use of this type of facility will be greatly extended. The control of simple main-line stations has been proved practical but there is still much to be done before the complete operation of pipe-line systems will be economical by remote control. Initial stations usually have extensive tankage and metering facilities requiring attendance for accounting reasons. The availability of highly simplified and reliable sequence control for tank-farm and main-unit operation makes local control feasible with minimum incremental personnel. A similar condition exists for stations at points where deliveries are made, common on lines transporting refined products. Here further complications are introduced by the necessity of timing the opening and closing of delivery valves according to the incidence of batch changes. However, it is often practical to handle deliveries and general maintenance work with one man engaged on an "on-call" basis, making remote control of the main-line pumps an economical possibility. One line is presently making an installation of supervisory control to a pumping station where deliveries will also be remotely controlled, involving the telemetering of batch changes and the positioning of a take-off valve from which refined products will flow through controlled manifold valves to the receiving tanks of the respective shippers.

There are so many variations in individual projects that the profitability of remote control and the control system arrangement must be determined by economic analysis for each specific case. With the engineering advances that are constantly being made these studies are continually becoming more rewarding.

Channels

The relatively long distances involved in the remote operation of most pipe-line pumping stations lend major importance to channel cost. Because of its economy in the use of channels, supervisory control must usually be chosen for remote operation. As previously discussed, supervisory-control systems are available which require only a single telegraphic-type channel.

The remote operation of pipe-line pumping stations usually requires continuous telemetered indications or recordings of quantities such as pressure and flow. Such telemetering is usually accomplished over telegraphic-type channels. The minimum number of channels for a particular installation is the sum of the supervisory control and telemetering channels. Commonly, additional channels will be required for voice communication and occasionally for other services, such as teletype, which require telegraphic-type channels. The necessary channels may be privately owned or leased.

Privately Owned Channels

Privately owned channels may be provided by telephone-type line wires or by microwave radio. Since pipe-line routes rarely conform to electric transmission lines, power-line carrier is not generally available as a channel means.

The telephone-line wires may be of open-type construction or may be conductors in a cable. Separate conductors may be used for each channel, or audio-frequency generating and receiving equipment may be employed to obtain 15 or more individual telegraphic-type channels over a single pair of wires. On open-construction line wires telephone-line carrier equipment can be applied. Such telephone-line carrier channels can be modulated to provide additional telegraphic-type channels. Regardless of the type of line-wire circuit employed, the channel cost is usually relatively high due to the distances involved.

Microwave radio is excellently suited to this service by virtue of its beam characteristic and its capacity for large numbers of channels. A single microwave beam may carry as many as 30 separate voice channels, each of which may be submultiplexed into as many as 15 telegraphic channels, if required. Through the use of repeater stations, microwave radio is applicable to pipe-line systems of any existing length, and is therefore suitable for systems employing centralized control as well as for 2-station links.[8,9]

Leased Channels

Voice channels can be leased and privately owned audio-tone generating and receiving equipment applied to obtain the required telegraphic-type channels. For minimum investment in equipment individual telegraphic-type channels can be leased in the number dictated by the supervisoy control and telemetering requirements.[10] It is immaterial whether these leased channels are obtained from metallic circuits, microwave-radio or telephone-line carrier, or any combination of these.

Choice of Channel

The choice of channel is a problem involving both economic and reliability considerations. It is not within the scope of this paper to discuss the economic and

Fig. 7 (left). This new booster station on a refined-products line operated unattended 56 miles downstream from the controlling panel shown in Fig. 6. Pumping unit is in building at right, separate from control building at left and microwave equipment in background

Fig. 8 (right). Pump room of remotely controlled station shown in Fig. 7. Pump is driven by 600-horsepower explosion-proof motor. Vibration-monitoring relay is visible on pump case

reliability factors governing the choice between leased and privately owned channels. As for privately owned channels, wire lines will usually prove more economical for short distances and relatively few channels, whereas microwave has an economic advantage for long distances or large numbers of channels. The importance of reliability in pipe-line operation justifies serious consideration of microwave even where its cost is relatively high.

The channel requirements reflect the pipe-line system arrangement as well as the telemetering and communication load to be carried. Because of the ability of supervisory control to provide a multiplicity of control and indication functions over a single telegraphic-type channel, the number of such functions does not ordinarily affect the choice of channel type.

Functions of Remote-Control System for Pipe-Line Service

TELEMETERING

It is generally desirable to telemeter the critical hydraulic quantities so that the operator may handle the remote station with the same type of intelligence that he has available for local-station control. The essential information for efficient operation generally can be gained from the pressures at initial pump suction, final pump discharge, and outbound line where the discharge is under throttling control. Except for changes incident to deliveries or injections, flow is constant throughout the line and generally, being known from instrumentation at other points, need not be telemetered from unattended booster stations.

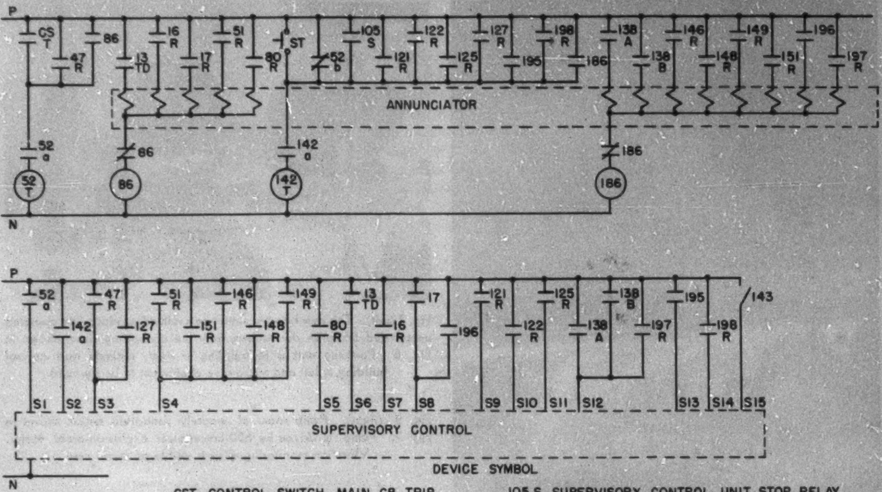
The essential pressures should preferably be telemetered continuously, although various forms of selective telemetering through supervisory control are available to economize in channels.

Two systems of telemetering are in common use. One is the pulse-duration type, wherein the measured quantity is transmitted in terms of repetitive pulses of variable duration.[11] This system transmits a definite number of pulses per minute (4, 12, or 30), and its speed of response is limited to this condition. Advantages of this system are its applicability to low-cost telegraphic channels, and the fact that the transmitter may be either directly actuated by the measured quantity such as pressure, or by the electric output of a suitable transducer.

Practically instantaneous response is attained by a system in which the measured quantity is delivered by a transducer to the transmitter as a direct voltage, 0 to 100 millivolts, which the transmitter then converts to a variable frequency, such as 15 to 35 cycles, that is transmitted over the telemetering channel to the receiver which reconverts it to d-c millivolts for indication or recording on an instrument suitably calibrated. This system requires a channel capable of handling relatively high keying speeds.[12,13] The choice between these two systems is one in which speed of response is weighed against investment in telemetering and channel equipment.

CONTROL AND INDICATION

The principal control and indication functions concern the starting and stopping of pump units, and the opening and closing of valves. Commonly each pump and its associated suction and discharge valves, and sometimes its ventilating fan, are operated in automatic sequence. Here the starting and stopping functions are reduced to the actuation of interposing relays for the unit-sequence control. In one form of supervisory control developed for pipe-line service this is accomplished at the controlling station by momentarily pressing a start and stop push button, panel mounted, frequently in a diagram showing the principal station piping; see Fig. 6. This one manipulation results in point selection, check-back and normally immediate operation of the equipment at the controlled station. Change in position of the controlled equipment is shown by associated light indication in an appropriate symbol on the diagram—green for



DEVICE SYMBOL

CST CONTROL SWITCH, MAIN CB TRIP
ST PUSHBUTTON, UNIT STOP
52 MAIN CIRCUIT BREAKER
86, 186 LOCKOUT RELAY, MANUALLY RESET
105 S SUPERVISORY CONTROL UNIT STOP RELAY
142 UNIT CIRCUIT BREAKER
143 UNIT REMOTE START CUTOFF
S1—S15 SUPERVISORY INDICATION POINTS

SYMBOLS FOR PROTECTIVE DEVICES AND PROTECTIVE AUXILIARY RELAYS ARE TABULATED SEPARATELY

Fig. 9. Schematic diagram of protection for an unattended single-unit pipe-line pumping station. For additional pumps corresponding unit devices would be installed, such as 205S, 221R, 222R, etc., for unit no. 2

down, amber for operating. In the case of reduced-voltage starters both lights show during the starting period. Valve positions are indicated by green or blue for closed, amber for open, with both lights burning for intermediate positions. In each case a central white light burns momentarily to indicate selection. Should selection fail to be verified, the white selection lamp will not light, and no further operation is performed until the push button is again depressed.

A remote-start cutoff supervision light at the controlling station informs the operator in the event that the system is blocked at the controlled station to prevent remote starting while maintenance work is being performed, or if the station is to be under attended operation for other reasons.

To keep the pipe internally clean it is usually necessary to run a scraper or "pig" through the line at periodic intervals. This scraper must be barred from the pumping stations. At attended stations valving is ordinarily provided either to allow the scraper to by-pass the station or to bring it into a "trap" on the upstream side, from which it is removed by the attendant and carried to the downstream side where it is reinserted through an outgoing trap. Recently considerable development work has been done on devices to signal the passage of a pig at a given point in the line. Such a detector may be installed just upstream of the station to cause shutdown upon arrival of the pig, the pig then proceeding through a full-opening station by-pass check valve, and through a second detector on the outgoing side to restart the station, restoring normal operation. Simultaneously the supervisory control system is actuated to provide an alarm and lamp indication at the controlling station while the scraper is in the station zone.

Protective functions are indicated by supervision lights at the appropriate points in the diagram—green for normal, red for trouble, white for selection. For all trouble points the red lamp is accompanied by an audible alarm which is silenced when a reset button is pressed. All trouble-lamp indications are maintained until the reset button is pushed, whereupon they disappear if the trouble has cleared; otherwise they continue until automatically extinguished upon clearance of the trouble. No alarm is given when a premeditated change occurs, such as transition of a valve from one position to another in a sequence-control system.

To permit sending the right maintenance personnel in case of trouble at the controlled station, and to provide the necessary intelligence for this with a minimum number of indications, several of the protective functions are usually grouped for a single indication. For example, a single lamp designated "electrical" may indicate overcurrent, short circuit, phase current unbalance, motor overtemperature, and incomplete sequence. Another, designated "mechanical," may include motor bearing, pump bearing, and vibration protection

Equipment Protection in Remotely Controlled Pipe-Line Stations

The hazardous nature of the medium being transported places special emphasis on the protective system in an unattended station. Failures must be minimized and the results should be controlled to the extent that an unattended station constitutes no greater hazard to life and property than the line itself. In this respect the uniformly good operating history of remotely controlled stations to date is justification of the careful attention given to this aspect of their design.

Fundamentally, the protective system should conform to the basic station layout, which is subject to considerable variation in individual projects. The most conservative engineering on new projects provides location of the control equipment completely apart from the pumping equipment, in separate buildings if the equipment is of indoor type. The booster station shown in Fig. 7 is an example of this practice. Outdoor equipment has obvious advantages in reducing construction cost and minimizing atmospheric hazards, but presents serious maintenance difficulties in bad weather. To date the majority of installations are indoor.

A considerable proportion of the remote-control installations are and will be at existing stations where the motor-starting equipment is installed in the same building with the pumping equipment. In attended stations using this arrangement good safety practice calls for installation of control, if of nonexplosion-proof construction, in a separate room maintained nonhazardous by pressurizing ventilation. This practice should not be relaxed, but rather, strengthened for unattended operation. Protective shutdown should be initiated either by loss of differential pressure between the control room and pump room, or by the approach to hazardous proportion of hydrocarbons in the control-room air as detected by a combustible-gas alarm. In modern attended stations the differential air-pressure relay is already quite common, usually being connected to provide alarm only, and it is a simple extension of this practice to incorporate it as a shutdown function in unattended stations.

Another feature requiring special consideration is that of shutdown in event of pump-seal failure. This may be accomplished by draining the seal leakage from the two ends of each pump to a common small sump equipped with an outflow orifice, or a small repump capable of handling normal leakage rates; then, upon occurrence of a seal failure the excessive flow into the sump causes a rise in liquid level, operating a float switch to shut down the unit. If the shutdown sequence includes automatic closure of the suction and discharge valves the unit is thereby isolated from the line and leakage will cease.

The main station sump should also be equipped with a float switch which will cause station shutdown on high level, reflecting an abnormal spillage of liquid in the station.

Mechanical failures in a main pump or its driving motor should initiate shutdown. In motors or pumps equipped with sleeve bearings this eventuality is covered by the use of bearing thermal relays which, however, to be effective, must be properly applied with the sensing element in close proximity, say 1/8 inch, to the journal surface. Thermal relays are not effective when they measure bearing-oil temperature because of the gradient and thermal capacity involved. Ball-bearing units are almost impossible to protect by thermal relays because there is usually no significant temperature rise before a bearing failure. However, ball bearings operate with a certain roughness upon incipient failure and the resulting vibration may be used to actuate a shutdown relay. A recently developed mechanical vibration-monitoring relay is applied to the pump case of the unit shown in Fig. 8. For effective protection such a relay must be very rigidly attached to the machine.

In addition to the protective functions just outlined as requisites for unattended stations, some of which are already coming into use in attended stations as well, the installation should include all those protective functions that would normally be provided in an attended station.

A recommended list of protective functions is given herewith, which includes the functions already discussed. All protective functions have individual or group indication at the controlling station. Some, designated *L*, cause lockout requiring resetting at the controlled station after the trouble is cleared. These have operation registered by a relay flag, or by a drop-flag annunciator, at the remote station. The symbols at the end of each item refer to Fig. 9.

STATION PROTECTIVE FUNCTIONS TRIP MAIN CIRCUIT BREAKER

L Ventilation failure or combustible-gas detection; used where nonexplosion-proof control is installed in pressurized control room in same building with pumping equipment. *13TD*
L Instrument air failure; where pneumatic pressure or flow controllers are used. *16R*
L High level in station sump. *17*
Power failure—time delay undervoltage protection with automatic reclosure of main breaker upon restoration of power with correct phase sequence. *47R*
L Overcurrent. *51R*
L Battery undervoltage; where d-c control is used. *80R*

UNIT PROTECTIVE FUNCTIONS SHUT DOWN INDIVIDUAL PUMP UNITS

Low suction pressure. *121R*
High pump case pressure, or high suction pressure. *122R*
High pressure in outbound line, or at final pump discharge. Note that in a multiunit station these functions may be arranged to trip units sequentially starting with the upstream unit. *125R*
Undervoltage time delay. *127R*
L Motor Bearing overtemperature. *138A, 138B*
L Phase current unbalance. *146R*
L Incomplete starting sequence. *148R*
L Overcurrent, running. *149R*
L Overcurrent, starting and short-circuit. *151R*
Pump case overtemperature. *195*
L Pump seal leakage. *196*
L Pump vibration; if ball bearings are used. *197R*
Scraper passage, with automatic restarting when passage is completed. *198R*

Fig. 9 shows, for a single-unit station, a typical arrangement of these protective functions in a control scheme for tripping

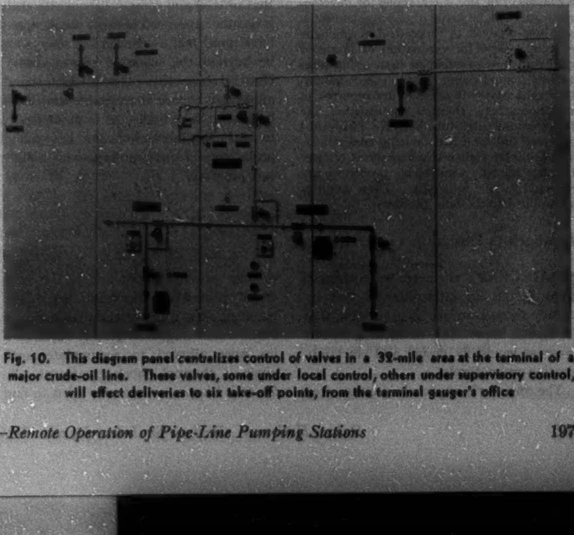


Fig. 10. This diagram panel centralizes control of valves in a 32-mile area at the terminal of a major crude-oil line. These valves, some under local control, others under supervisory control, will effect deliveries to six take-off points, from the terminal gauger's office

and indication at the controlling station, and, where specified, for lockout with annunciation at the controlled station. It is important to note that all protective functions act directly upon the concerned apparatus and through circuits which are entirely within the controlled station. The supervisory control system is not a link in the protective system but only serves to report and identify at the controlling station the operation of a protective function at the controlled station.

Conclusion

With improvements in basic station layout and the development of more complete and more reliable protective systems for automatic shutdown in event of abnormal operation or equipment failure, it has proved practical to operate pipeline pumping stations remotely with reduced station attendance. Recently, the completely unattended booster station has become an operating actuality.

Many systems for remote control have been applied. In comparing systems which perform the desired functions the most important considerations are reliability and cost. In making economic comparisons both initial and operating costs must be considered. The initial costs include the terminal equipment at the controlling and controlled stations and the necessary channel equipment. The operating costs include channel rental if leased channels are employed, plus maintenance, and cost of attended operation at such times as the remote control system is out of service, as well as cost of throughput loss during down times chargeable to the remote control system.

The hazardous nature of the medium transported makes reliability the most important single consideration. Reliability is increased by the use of adequate protective equipment at the controlled station. Reliability of remote operation is enhanced by using supervisory control equipment which is essentially trouble-free and provides complete supervision of the operating channel.

References

1. Distant Pump Station Successfully Worked by Remote Control, **L. G. E. Bignell.** *Oil and Gas Journal*, Tulsa, Okla., Dec. 1931, pp. 18, 94.

2. American Standard Definitions of Electrical terms. *AIEE*, 1941.

3. A New 2-Signal Synchronous Supervisory Control System, **W. A. Derr.** *AIEE Transactions*, vol. 68, pt. II, June, 1949, pp. 857–63.

4. Development of a Two-Wire Supervisory Control System, **R. J. Wensley, W. M. Donovan.** *AIEE Transactions*, vol. 28, Oct. 1930, pp. 1339–43.

5. By a Flick of the Finger (Supervisory Control), **W. A. Derr.** *Westinghouse Engineer*, East Pittsburgh, Pa., Nov. 1949, pp. 162–67.

6. Space Code Selector Supervisory System, **E. F. Forrest, P. W. Schirmer.** *AIEE Transactions*, vol. 68, pt. II, Sept. 1949, pp. 1312–16.

7. A Self-Checking System of Supervisory Control, **M. E. Reagan.** *AIEE Transactions (Electrical Engineering)*, vol. 47, Oct. 1938, pp. 600–05.

8. Microwave Communication, **N. B. Tharp.** *Westinghouse Engineer*, East Pittsburgh, Pa., Nov. 1952, pp. 198–203.

9. Microwaves in the Electric-Power Field, **W. A. Derr, D. F. Burnside, H. W. Lensner.** *AIEE Transactions*, vol. 71, pt. III, July 1952, pp. 910–17.

10. Leased Channels for Metering and Control, **H. H. Joyner.** *Electric Light and Power*, Chicago, Ill., Dec. 1953, pp. 100–06.

11. Telemetering and Remote Control, **W. E. Rufleth.** *Petroleum Engineer*, Dallas, Tex., Oct. 1952.

12. An Electric Telemetering System, **Carl Oman.** *Radio and Television News*, Radio-Electronic Engineering Edition, Chicago, Ill., Oct. 1952, pp. 14–16.

13. A High-Speed Telemetering System with Automatic Calibration, **W. E. Phillips.** *AIEE Transactions*, vol. 70, pt. II, June 1951, pp. 1256–60.

◆

No Discussion

Evaluation of Designs for Intermittently Heated Surfaces

THOMAS M. DAHM
NONMEMBER AIEE

RAYMOND A. HOLLOWAY
ASSOCIATE MEMBER AIEE

Synopsis: In the early stages of development, use can be made of thermal transmission line analysis of theoretical or analogue computer type for evaluation of proposed configurations of such an item as an electrically heated deicer boot. Direct laboratory evaluations of samples are reported in which synthetic thermal transmission line terminations substitute both for theoretical analysis and for flight or icing tunnel test. The findings of theoretical analyses of electrically heated windshield problems are reported in the Appendix. These findings are applied to the boot problem.

Approach to Design

If AIRCRAFT engineers were assigned to design an intermittently but intensely heated surface, it is possible that the great importance of making visible progress would fix design and start fabrication prematurely.

The engineer's approach is quite often that of trial and error. Frequently there is no other approach because the idea may have novel features or may in some ways be beyond the scope of known theory. But there are other cases in which sound, applicable theory is available and could be used at very little cost to guide design to producible fabrications that have a high probability of successfully passing a flight test program.

Propeller Deicer Boots

A well-known example of the supposedly practical approach is the electrically heated propeller deicer boot. The excellent idea was presented that power requirement might be greatly reduced if ice were allowed to accumulate until given a thermal shock treatment. So far as is known, no one ever tried to find out, for example, how hot the heating wires become. Perhaps no one realized that excessive heat storage at the source slows action. Highly competent rubber technology was applied, but thermally and electrically the development of the boots was unpremeditated. It is understood that one manufacturer added a metal outside surface with the idea of sacrificing some speed in deicing to gain considerably longer boot life. No one seems to have been unduly curious as to why the addition of thermal capacity speeded up the boot, contrary to expectations.

Leading-Edge Electrical Deicing

When electrical deicing of leading edges was attempted, no basic development took place. All attention was directed to solving the production difficulties of much larger boots of the same patterns as

---

Paper 54-308, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted August 12, 1953; made available for printing June 10, 1954.

Thomas M. Dahm and Raymond A. Holloway are with the Lockheed Aircraft Corporation, Burbank, Calif.

before. No one apparently asked whether certain components required to withstand centrifugal forces on propellers might now be reduced. Apparently no one investigated the placement of poorly conducting material, such as heavy thread containing air. Integration of the radially directed temperature gradient from a hot wire shows that 0.005-inch radial thickness of such material in contact with the wire and surrounded by rubber of higher thermal conductivity produces much higher source temperature and consequent heat storage at the source than a much larger amount of the thread placed as a concentric layer, also 0.005 inch thick, at some distance in rubber from the wire. It is particularly important for either transient or steady-state operation to place material of low thermal conductivity elsewhere than at points where geometrical factors produce a high density of heat flow.

Thermal Analyzer

At about the time that intermittent electrical deicing of leading edges began to be seriously considered, Tribus[1] used a thermal analyzer, intended to represent both a wired deicer boot and some heat loss conditions representative of icing situations. These heat losses were represented by several ingenious nonlinear circuits.

Fineness of structure for representation of distributed thermal resistance and capacity by lumped constants was attempted at points thermally distant from the heat source, where it was found unimportant. But the boot itself forward of the heat source was represented by a single resistor center-tapped by a condenser. Electrical engineers familiar with long line analysis would expect that outside surface temperature changes soon after start or stop of heating current will not be well represented by this lumped circuit.

An important result of this evaluation of intermittent heating by other means than trial-and-error is the emphasis on the potential usefulness of heating at high intensity for very short time intervals for periodically removing ice accumulated during power-off intervals. It has long been recognized that electrical power requirement for steady-state anti-icing would be much too large because it would be necessary to evaporate completely the droplet catch before the water could run back and refreeze behind the heated area on the leading edge of a wing. But intermittent heating at high intensity will also not be effective unless its shortness of input time is associated with shortness of cooling time. The latter is very dependent on thermal transmission line characteristics.

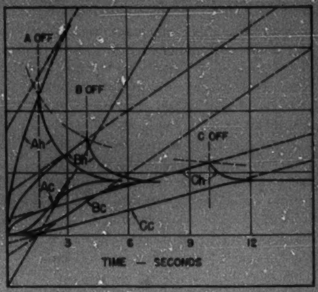
It is obvious that electrical heating can be stopped and started very suddenly. It is not so obvious that a square-wave heat pulse can be watered down to a long, low hump of temperature at a short distance from the heat source. The duration of this hump determines the time required for the boot to cool after the heating current stops. This time determines how much run-back and refreezing beyond the aft edge of the heated area are to be expected before freezing begins at this area during the off phase of the heating cycle. The general nature of source and of cooler surface temperature curves is indicated by curves *Ah* and *Ac* of Fig. 1, except that these are drawn for a condition of no heat loss at the surfaces.

Fig. 1. Heat pulse into 3-second flat plate, unilaterally heated and insulated

Ah, Bh, Ch—Hot-surface temperature versus time

Ac, Bc, Cc—Cold-surface temperature versus time

	A	B	C
On time seconds	1.6	40	10.0
Heat flow, But/hr ft^2	6.25Q	2.5Q	Q
Heat pulse But/ft^2	Q/360	Q/360	Q/360



Icing Phenomena Point of View

It is to be expected that engineers who have specialized in weather and icing phenomena will devote much of their attention and effort to icing flight test or to work with tunnels for icing research. Their attitude toward the device that supplies the thermal energy may be indicated by the question "What temperature is necessary on the boot surface to effect deicing?" This question may appear uncalled-for but practical experience shows that commercially available wired de-icer boots have interesting surface temperature patterns, with temperature much higher than 32 degrees Fahrenheit (F) at some points, while a considerable fraction of area is still below 32 F.

Transmission Line Point of View

It is much easier for an electrical engineer to appreciate the significance of the fact that the time required for a thermal change at the heat source to reach the surface of a thin boot is comparable to the time it takes an Atlantic cable to produce a strong signal at its far end. With this background, the electrical engineer will be inclined to investigate the boot as a thermal transmission line. He will also view weather phenomena as constituting thermal loads terminating the transmission line.

Related Electrical and Thermal Transmission Lines

The Thury type of transmission line developed and operated for nearly 50 years in France[2] and an American line built for investigation of communication methods[3] are basically constant-current d-c lines. Transients associated with starting ideal infinite lines were partially investigated by Heaviside. A constant direct current is suddenly supplied to the common home end of two infinite lines having different resistance and capacitance per mile. The proportions in which current divides between the lines, the history of home end rise of voltage, and the instantaneous value of current at distant points were determined. Integration of the Heaviside series expression for current gives the changing pattern of rise of voltage.

These Heaviside results are shown in the thermal terms in the Appendix. The result of the integration is equation 8, which can be arranged as a product of three factors contributing to temperature rise:

1. Heat input rate Q.
2. Home end thermal impedance $Z = (4\theta/\pi k c \rho)^{1/2}$.
3. A nondimensional factor $f(y)$ that decreases from unity as y increases. The quantity $y = x^2/4\alpha\theta$ is large when distance x is large or when time θ is small.

It is very important to note that this third factor is not changed when distance from the source is changed, provided that a compensatory time change is made. But if distance is increased in a ratio r, time must be increased in a ratio r^2. While the whole story is not as simple as this, the surprising time lag between a seasonal temperature peak and its appearance at depth in the earth shows that large time change must be traded for small distance change. A boot designer would be wise to avoid a configuration in which heat is readily conducted forward and equally readily conducted aftward toward a heat barrier which limits escape to structure. This heat that starts aftward needs a long time to get to the forward surface during the cooling phase. When rapid cooling after a burst of heat is required to limit run-back and refreezing, the designer has two options:

1. If adequate electric power is available, he places behind the internal heating element thin plies of material which conducts as well thermally as the plies forward of the heat source, and he establishes good thermal contact between these plies and the airplane structure to insure that any heat which starts aftward continues in that direction.
2. With limited power available, he places on the aft side material which contains as much air as possible as close as possible to the heating element, to limit the amount of the heat which starts aftward but finally comes out forward.

Bracketing Thermal Loads

The infinite line solution can be modified in various ways for application to practical problems involving various types of load terminations of long lines. Among these terminal loads are two that bracket the range of possible loads. It is useful to open-circuit the long line and start the constant current into the home end; also to ground the distant point and start the current.

In the thermal case, open-circuit condition is closely approached when a section of the deicer boot is roofed over by ice with which it is not in good thermal contact. Short-circuit condition is surprisingly well simulated by an extremely cold ice film that will not melt because, at its weather surface, heat energy is being converted into energy of change of state. The weather surface is nearly locked down in temperature because is is not much warmer than the passing air when sublimation converts ice to vapor. Under these conditions, the ice is a transmission line in series with the boot. All icing weather phenomena lie between these bracketing conditions. Inferior designs can often be rejected at an early stage with the aid of this kind of analysis.

It is not implied that all of the problems can be solved by analysis. For example, a commercially available deicer boot was seen in a laboratory icing test to have melting and freezing patterns which could have very adverse effects on ice removal. These patterns were also made visible in color on small samples by using an aqueous solution containing a substance that fluoresces only weakly when the solution is frozen. Application of ideas gained from transient heat flow analysis determined the causes of these patterns sufficiently well to contraindicate a difficult program for determination of the fine structure of boot surface temperature.

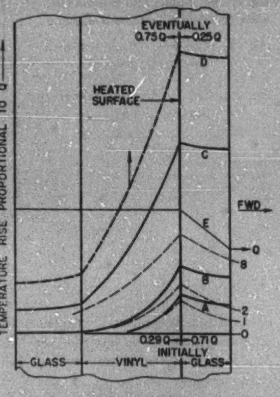


Fig. 2. Temperature rise versus distance from heat source, internally heated insulated bird-proof windshield

1, 2, 8—Semi-infinite medium distributions after 1, 2, and 8 minutes

A, B, C—Insulated panel distributions after 1, 2, and 8 minutes

D—Steady-state of rise distribution curve moves up at $\theta/\Sigma c\rho\ x$ F/hr

E—Steady-state flight distribution

Application of a Heaviside Solution to a Terminated Line

Several years ago the Heaviside infinite line solution described in the Appendix was applied to the problem of an electrically heated windshield for the purpose of determining the temperature distribution in it during ground warm-up, a condition that approximates the zero heat loss of open circuit.

The 3/16-inch forward glass, (with the bussed conducting coating on its aft surface) was bonded to a thick bird-proof vinyl ply between it and a 1/4-inch aft glass. Because the plastic conducts heat about one-fifth as well as glass, the temperature pattern after some minutes of design power application is rather startling. In particular, how much hotter is the source than the forward surface? How long do the thin forward glass and the thick plastic ply behave like semi-infinite media? How does the heat divide between them initially, and does the ratio change? How long does it take to reach a steady state of rise condition throughout, with the aft glass far behind in temperature? What sort of temperature gradient exists in the plastic at the sensing element location or how much does this wire grid or thermistor lag the heat source?

Some of these questions can be answered on the basis of the semi-infinite line solutions. Experience with the thermal Ohm's law commonly used in analysis of steady-state heat transmission would lead to the expectation that if vinyl conducts heat one fifth as well as glass, one sixth of the heat will flow into vinyl and five sixths into glass as soon as the conducting coating between these materials is energized. Equations 2 and 3 show that the vinyl will take a higher proportion than one sixth because acceptance of heat flow is determined by $(kc\rho)^{1/2}$ which is the geometric average of two properties of a material, which take heat away from the source, thermal conductivity k and volumetric heat capacity $c\rho$. In the analogous case of the constant-current transmission line, two combinations of constants are important, $(c/r)^{1/2}$, analogous to $(kc\rho)^{1/2}$, and cr the time constant analogous to the reciprocal of diffusivity α, where $\alpha = k/c\rho$. The initial division of flow between the two media on the basis of estimated physical constants is shown at the bottom of Fig. 2, 71 per cent into glass and 29 per cent into vinyl plastic.

At least for a short time after heat input starts, home end temperature increases with the square root of time. Curves 1, 2, and 8 of Fig. 2 show the temperature distributions that would exist after 1, 2, and 8 minutes if both media extended indefinitely. It is evident that the vinyl continues for a long time to act as an infinite medium, both because it is thicker than the forward glass and because it does not conduct heat as well.

The curves A, B, and C represent 1-, 2-, and 8-minute temperature response of this system of three thermal transmission lines having two open-circuited terminations. They result from summation of effects of a system of image sources and sinks described in the Appendix. Curve D, moving upward as indicated by the vertical arrow, shows the eventual steadily rising temperature distribution that is nearly reached in about 18 minutes. About three quarters of the heat is then flowing into the vinyl because about three quarters of panel heat capacity is behind the coating. Curve *E* represents an eventual steady-state delivery of heat forward as determined by an assumed flight heat transfer coefficient. Its horizontal section implies that source temperature is the same as cockpit air temperature. Even when this is not the case, heat transfer through the body of the panel produces low gradients because transfer at the cockpit surface is held down by the usually low value of the transfer coefficient at this surface.

For time intervals up to 1 minute after the end of a heating cycle, the drop in temperature near the source is well represented by superposing an inverted infinite medium curve on the steady-state curve. Consideration of transfer impedance (ratio of temperature rise to heat flow at any instant at a given distance from the heat source) shows that the infinite medium could take heat away from the location of the actual weather surface much faster than could a flight heat transfer coefficient during an early, short-time interval. It follows that both source and weather surface temperatures are well approximated at 1 minute by inverting and superposing curve *A* instead of curve 1. This kind of information is very important to design of a temperature control system by which time-average power is adjusted to need by on and off cycling.

It is visually evident that inversion of an infinite medium curve in the vinyl and superposition on the steady-state curve evaluates the temperature difference between the source and a temperature-sensing element in the vinyl at times considerably longer after heating is stopped than can be used on the glass side.

Computation was not attempted of effects of "grounding" the forward surface, or of a resistance load equal to the reciprocal of a transfer coefficient. For this purpose analogue or other types of computers should be used. These give good numerical answers, even though they do little to establish functional relationships or to improve intuition. Frequently, the analytical solution provides very reliable means of determining whether the machine is working correctly. Many flight conditions have been evaluated with at least two types of computer.

Related Thermal Transient Analysis

Jakob[4] presents useful information on analysis of transient heat flow problems, including cases of sudden start of constant heat flow and cases of sudden change of surface temperature. The solutions do not involve Heaviside terminology.

Industrial use is made of a Heaviside solution developed by engineers of the Boeing Airplane Company. A major difficulty with electrically heated windshields is that coatings of closely controlled surface resistivity cannot be produced. Structural and aerodynamic considerations often force panel shape which does not sufficiently approximate rectangularity. It is necessary to determine before lamination whether a glass on which the conductive coating has been placed will meet power-constant requirements. The analysis predicts that after about a quarter of a minute the hot and cold sides of the suddenly energized glass begin to rise in temperature at the same constant rate, and with a difference in temperature that is proportional to the heat output at the measurement point. This solution is generalized in the Appendix and forms the basis for Figs. 1 and 3. The latter shows the way in which a flat plate unilaterally heated would develop time patterns of surface temperature when heated at different intensities for such time intervals as result in supplying the plate with identical quantities of heat. These are discussed in the Appendix as qualitatively applicable to wired deicer boots.

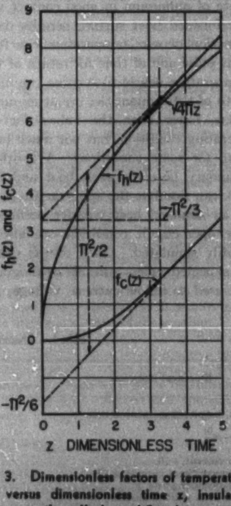


Fig. 3. Dimensionless factors of temperature rise versus dimensionless time z, insulated unilaterally heated flat plate

$f_h(z)$—Hot-surface factor
$f_c(z)$—Cold-surface factor
$(4\pi z)^{1/2}$—Hot-surface parabolic factor as for semi-infinite medium
$z = \pi^2 \alpha\theta / x_1^2$

The Armour Research Foundation is applying Laplace transform methods and modifications of previously published approximate methods as one phase of a development contract on internally heated windshields.

Plastics and High Mach Number

The ability of plastic items such as canopies to withstand high temperature that occurs during a burst of speed is a matter of grave concern. At a speed slightly greater than 1.5 times the speed of sound, the weather surface of the plastic may rise 200 F. Laboratory tests show that it is possible to raise the surface temperature of a plastic very fast indeed. Consequently, a theoretical solution that assumes instantaneous change of surface temperature is almost as closely followed in practice as in the case of instantaneous start of a constant heat flow when the heat is produced electrically in a very thin coating or is produced by radiation from a source very much hotter than the irradiated surface. Rise in temperature during a burst of speed not too many seconds long is obtained by substituting temperature for heat flow in equation 2, to give

$$t = t_1\left(1 - 2\pi^{-1/2}\int_0^u e^{-u^2}du\right)$$
$$= t_1\left(1 - \mathit{erf}\, x(4\alpha\theta)^{-1/2}\right)$$

The internal temperature is t at time θ and at various distances x from the surface that is suddenly raised t_1 degrees by the burst of speed It can be computed from *erf u* tables.[4,5] A suitable value of diffusivity $\alpha = k/c\rho$ is not too far from $0.12/27.5 = 0.0044$ ft²/hr.

If the burst of speed is of brief duration and if the canopy is not too thin, the internal temperature distribution will be approximated by computation of this infinite line case. A better approximation will be obtained by taking the inner surface rise in temperature as double the value computed on the infinite line basis. Such a system of images as was used for the constant heat input case is not quite applicable to this case, for boundary condition reasons. If necessary, the problem may be completely solved by Fourier analysis methods, as in chap. 14, paragraphs 5 and 6 of reference 4.

The possible value of such a solution is to determine how thick present thermoplastic canopies must be to have enough thickness of cool plastic to withstand pressurization at the end of the burst of speed and during the interval required for penetration of heat stored near the hot outside surface.

Laboratory Application of Principles of Transient Analysis

Solution of the way in which widely spaced heating wires deliver a heat pulse to a boot surface and aftward to supporting structure may not be practicable by analysis. An analogue computer solution will be no better than the network assumed to represent the configuration. Neither will appear to be worth the effort if one is convinced that this type of configuration must constitute a slow line. The small increase in the speed of heat release resulting experimentally from reduction of the forward ply of a deicing boot to half thickness shows that delay in delivery of heat to the boot surface is very largely the result of excessive heat storage in and near small heated wires that are too widely spaced. Nevertheless, it has been found possible to prove quite simply which of two types of configuration can deliver heat most quickly. If fore and aft aluminum plates are given good thermal contact to a section of a boot, the temperature histories of these plates through a heating cycle tell most of the story without either the expense or uncertainties of icing flight.

It is quite possible to produce synthetic ice layers and boundary layers. Rubber has too low a thermal conductivity to represent both the resistance and capacity of an ice layer. But combination of rubber with metal foil between the boot and a massive heat sink of metal that represents all outdoors can simulate (up to melting point of ice) a layer of very cold ice converting heat to sublimation energy. It might prove that the ice cannot receive the heat fast enough to reach melting point at the boot surface. The obvious procedure here is to supply more heat. A less obvious procedure, but a better one, is to deliver the heat to the ice at a higher rate. The way to do this is to change the time constant of the transmission line. Sometimes the best way to do this is to open out some thermal bottlenecks very near the heat source.

Radial heat flow from widely spaced wires is such a bottleneck. But if the 10 strands of, say, a no. 26 wire were woven at close spacing into a flat thin ribbon, the wire surface out of which the heat flows into an insulating material is increased in about the ratio 3. The temperature rise of the wire during an on interval will be reduced, though not in the ratio 3. The spread in boot surface temperature will be reduced in a very large ratio. It is to be expected that the heat input per cycle necessary to melt ice at a sufficient percentage of ice contact area will be reduced by the amount of heat wasted in melting a large amount of ice over each of the widely spaced larger stranded wires. This reduced heat can be put into the finer wires in a very much shorter time without overheating.

It may well be that practicable wire patterns will not come close enough to a plane heat source to warrant full confidence in either a mathematical analysis or an analogue computer analysis. But a large variety of configurations can be fabricated as small samples and can be compared in the laboratory at the bracketing conditions equivalent to open or grounded electrical transmission lines.

A very limited amount of work has been done with a small boot sample in which wire with a substantial temperature coefficient was used. A significant item is the time it takes the wire temperature to drop nearly to an asymptote after a heat pulse into a thermally insulated boot. The first test was a pulse with decreasing power because of resistance change of the heated wire. Later tests are planned in which substantially constant power will be obtained by including a constant resistance ballast that matches varying wire resistance at some point. A very thick plate of aluminum in good contact with the forward boot surface permits determination of wire temperature rise for a heat pulse and of time for return of wire temperature nearly to starting condition. Both of these quantities are at minimum value, lower than the values for any operating condition, but not much lower than for some conditions. A synthetic boundary layer between boot and heat sink can simulate the approximate effects of an assumed heat transfer coefficient. The open-circuit condition can also be readily simulated. Good results cannot be expected unless suitable oscillographs are used to follow current, voltage, and the temperatures of thermocouples, some of which must be very thin.

There seems to be no great difficulty in hand-looming a variety of wiring configurations or in assembling them to fore and aft plies of different thickness or thermal conductivity. They can then be compared at the bracketing conditions. It is believed that this laboratory mockup of an analytical process would give useful results in cases too difficult for theoretical or machine analysis.

Appendix. Parallel Heat Flow into Semi-Infinite Media

The partial differential equation applicable either to parallel heat flow q or to rise in temperature t, as determined by the boundary conditions, is

$$\frac{\partial^2 q}{\partial x^2}=(c\rho/k)\frac{\partial q}{\partial \theta} \text{ or } \frac{\partial^2 t}{\partial x^2}=(c\rho/k)\frac{\partial t}{\partial \theta} \qquad (1)$$

Table I identifies the thermal quantities and lists the corresponding electrical quantities for the analagous electrical transmission line problem. Translated into thermal terms, Heaviside's solution (for the case of switching a constant flow of direct current into the common home end of two infinite transmission lines of differing resistance per mile and capacitance per mile) gives the heat flow in the medium with subscript 1 at any given time θ after heat input Q into the home end starts and at any distance x_1 from the heat source

$$q_1=Q[(k_1c_1\rho_1)^{1/2}/((k_1c_1\rho_1)^{1/2}+(k_2c_2\rho_2)^{1/2})][1-erf\, x_1(4\alpha\theta)^{-1/2}] \qquad (2)$$

Tables are available[4,5] for evaluation of Gauss' error function, the integral under the normal probability curve

$$erf\ u=2\pi^{-1/2}\int_0^u e^{-u^2}\,du$$

$$u=x_1(4\alpha\theta)^{-1/2}$$

Values of *erf u* may also be computed from the series resulting from term by term integration of

$$e^{-u^2}=1-u^2/1!+u^4/2!-u^6/3!+\ldots$$

Table I. Electrical and Thermal Quantities

Electrical Notation	Thermal Notation
Semi-infinite transmission line	Semi-infinite conducting medium
$\frac{\partial^2 i}{\partial x^2}=cr\frac{\partial i}{\partial t}$	$\frac{\partial^2 q}{\partial x^2}=(1/\alpha)\frac{\partial q}{\partial \theta}$
Zero current, $t<0$	Zero heat flow, $\theta<0$
Constant direct current, $t>0$	Constant heat input, $\theta>0$
Distance along line, x	Distance from heat source, x ft
Time after start of current, t	Elapsed heating time, θ hrs
Voltage rise at $(0, t)$, e_0	Temperature rise at $(0, \theta)$, t_0 F
Voltage rise at (x, t), e	Temperature rise at (x, θ), t F
Constant current at $(0, t)$, I	Constant heat input at $(0, \theta)$, Q Btu/(hr)(ft^2)
Current at (x, t), i	Heat flow at (x, θ), q Btu/(hr)(ft^2)
Reciprocal, resistance/length, l/r	Thermal conductivity, k Btu/(hr)(ft^2)(F/ft)
Capacitance/length, c	Volumetric specific heat, $c\rho$ Btu/(ft^3)(F)
Admittance factor, $(c/r)^{1/2}$	Admittance factor $(kc\rho)^{1/2}$, Btu/(hr$^{1/2}$)(ft^2)(F)
Reciprocal of time constant, $1/cr$	Diffusivity $(k/c\rho)$, α ft^2/hr

The values of *erf u* are somewhat larger than *u* up to about $u=0.6$ At $u=3.0$ *erf u* differs negligibly from the convergence value 1.0, but the series computation becomes laborious for larger *u*.

1. The form of equation 2 shows that Q divides permanently between the media in the ratio

$$m=Q_1/Q_2=(k_1c_1\rho_1/k_2c_2\rho_2)^{1/2} \tag{3}$$

2. Within the media, paired points x_1 and x_2 can be found at which $q_1/Q_1=q_2/Q_2$ at any given time. These points are related by

$$x_1/x_2=(\alpha_1/\alpha_2)^{1/2} \tag{4}$$

3. Since $q=-k\partial t/\partial x$, the slopes of the temperature curves at these paired points are in inverse ratio to the values of $\alpha^{1/2}$.

4. The temperature increments at any given time θ are identical at these paired points.

The home end rise of potential found by Heaviside becomes

$$t_0=Q_1(4\theta/\pi k_1c_1\rho_1)^{1/2}=Q_2(4\theta/\pi k_2c_2\rho_2)^{1/2} \tag{5}$$

By analogy with the electrical case the home end thermal impedance is, for medium 1

$$Z_1=(4\theta/\pi k_1c_1\rho_1)^{1/2} \tag{6}$$

When the flow into medium 1 is written Q_1, equation 2 may be rewritten for a single medium in the series form in which Heaviside obtained it, and from that

$$\begin{aligned}\frac{\partial t}{\partial x}=-q_1/k_1=-Q_1/k_1+(2Q_1/k_1\pi^{1/2})\times\\ [(x^2/4\alpha\theta)^{1/2}-(1/3)(x^2/4\alpha\theta)^{3/2}+\\ (1/5\times2!)(\quad)^{5/2}-\\ (1/7\times3!)(\quad)^{7/2}+\ldots]\end{aligned}$$

$$\begin{aligned}\frac{\partial t}{dx}=-Q_1/k_1+Q_1(4\theta/\pi k_1c_1\rho_1)^{1/2}\times\\ [2x/4\alpha\theta-(2/3)x^3/(4\alpha\theta)^2+\\ (2/5\times2!)x^5/4\alpha\theta)^3-\ldots]\end{aligned} \tag{7}$$

Term by term integration with respect to x from $x=0$ where $t=t_0$ to $x=x$, where $t=t$ gives, dropping subscripts, and simplifying coefficients

$$\begin{aligned}t=QZf(y)=Q(4\theta/\pi kc\rho)^{1/2}[-(\pi y)^{1/2}+\\ 1+y-y^2/3\times2!+y^3/5\times3!-\ldots\\ (-1)^{n+1}y^n/(2n-1)n!]\end{aligned} \tag{8}$$

where $y=x^2/4\alpha\theta$, the dimensionless power series variable.

Note that the first series term in y is preceded by $+1$ to give the home end value t_0 for $y=0$. The term $-(\pi y)^{1/2}$ when multiplied by the preceding factors is the integrated negative slope term that characterizes steady-state heat conduction, $-Qx/k$. The ratio of the nth series term to the preceding term is $-(2n-3)y/(2n-1)n$.

At the home end, $x=0$, the temperature rise is proportional to Q and to the square root of time, $\theta^{1/2}$. Initially, $dt/d\theta$ is infinite at the home end for an infinitely thin plane heat source.

Equation 7 shows that the entering gradient remains permanently $-Q/k$. When y becomes smaller with increasing θ, the entering slope tends to prolong itself into the medium so that the portion of the curve near $x=0$ approaches the constant slope steady-state conduction condition.

On the basis of equations 3 and 4 it follows that any part of a flow in medium 1 of high $k_1c_1\rho_1$ value that crosses into medium 2 of low $k_2c_2\rho_2$ value will have a temperature-raising effect in medium 2 that is m times as great as is produced by this same flow in medium 1 if (as also at home end)

$$m=(k_1c_1\rho_1/k_2c_2\rho_2)^{1/2} \tag{9}$$

Application to Windshield Problem

The third factor $f(y)$ in equation 8 can be evaluated in terms of y as it falls from unit value at $y=0$ toward negligible values when y is large either because of large x or small θ. Six-place logarithm tables become inadequate when y becomes large because the first series terms become large and alternate in sign. The final result is a very small difference between large sums. Present-day machines can carry the result out far enough, or trends of successive ratios of values of $f(y)$ can be extrapolated to get approximate low values of $f(y)$ at large values of y. Temperatures at a given time after start were calculated for equispaced values of x in glass, using $k=0.57$ Btu/(hr)(ft²)(F/ft) and $c\rho=32.0$ Btu/ft³F. Vinyl values used were $k=0.11$ Btu/(hr)(ft²)(F/ft) and $c\rho=27.5$ Btu/ft³F.

Image Flows

The infinite line solutions for the two media with common home end can be used to obtain the temperature rise of an insulated windshield by introducing the concept of heat "reflection," total at insulated surfaces and partial at interfaces between two media. The particular nature of this reflection is arrived at, as in many other cases, by introducing fictitious mirror image sources. In this case all image flows start at the same moment as the actual heat source.

If the forward glass were assumed to continue beyond its actual forward surface and if a fictitious source of strength Q_1 (equal to the actual source that produces forward-flowing heat) were to drive heat aftward, the two flows would produce zero heat transfer at the physical forward surface equidistant from both sources. But the temperature rise that would be produced at this point at a given time by the image flow would be superposed on the rise produced by the infinite medium flow from the real source. The early effect of the first image source is to double the rise in temperature that would occur at 3/16 inch from the heat source in a semi-infinite medium.

When this first aftward image flow q_i is evaluated at the same time θ at actual source position (glass to vinyl interface 3/8 inch from image position) it must be considered as dividing into reflected and transmitted flows q_r and q_t subject to equation 9.

$$q_i+q_r=mq_t \tag{10}$$

q_i flowing aftward in glass, q_t aftward in vinyl, q_r forward in glass. This is equivalent to requiring all three flows to come from sources of such strength and at such distances in glass for q_i and q_r, and in vinyl for q_t, as to maintain continuity of temperature across the interface.

Conservation of energy gives

$$q_i-q_r=q_t \tag{11}$$

solving

$$q_t=[2/(m+1)]q_i \tag{12}$$

$$q_r=[(m-1)/(m+1)]q_i \tag{13}$$

These multipliers of q_i designate strengths of the second and third image flows which furnish components of temperature rise. The third flow q_r has its source located in glass in the obvious position of a mirror image of the image that produces q_i. But the second flow q_t considered as coming from a source in vinyl, is subject to the x distance shrinkage of equation 4, just as in the case of the actual initial Q_2 flow into the vinyl. The temperature-raising effects of these flows are scaled with suitable multipliers from the $f(y)$ curve of equation 8, as computed for a fixed value of θ and a series of values of x. (The $f(y)$ curve is more convenient when plotted with y values as abscissas, rather than x values). The forward flow q_r is subject to a total reflection at the actual forward glass surface, and so on, to negligible increment of temperature.

Computations at early value of time give negligible temperature rise at the aft surface of the thick vinyl ply, infinite medium behavior. For computation at a later time, aftward incidence and transmission at the aft vinyl-glass interface make forward reflection q_r negative because m in equation 13 is less than 1 for this sequence of media. Glass drawing heat out of vinyl is represented by a sink image in imaginary vinyl aft of the interface that produces temperature decrements in the actual vinyl forward of this interface. In the case of 1 or 2 minutes of heating, only a few sources or sinks contribute significantly. A computation for 8 minutes was rather laborious.

Computation accuracy was indicated by the fact that all points fitted smooth curves and by a check of the integrated value of assumed heat input against the sum of the three heat quantities obtained by multiplying the area under each temperature curve by the value of $c\rho$, the volumetric specific heat of each material.

A laboratory check was also made. A windshield panel was brought to a steady state of low temperature by circulation of a cold liquid over its forward surface. Power was applied as soon as possible after draining off the cold liquid. Instrumentation had been assembled to measure the course of panel surface temperatures and to measure the resistance of a thermistor temperature-sensing element inside the vinyl ply. The temperature histories were in good agreement with predicted surface temperature for about 5 minutes. After that length of time, heat loss from the forward surface reduced the rate of rise of forward surface temperature below that proper to the thermally insulated case.

Fig. 2 indicates schematically the relative rise in temperature at various points in a bird-proof insulated panel after given heating intervals of 1, 2, and 8 minutes, curves *A*, *B*, and *C*. It shows the shape of the infinite medium solutions at corresponding times, the shape of the eventual steady-state pattern *E* for any given heat transfer coefficient in flight, and an asymptotically approached steady state of rise pattern *D* for the insulated case. In this steadily rising pattern the temperature distribution curves are segments of parabolas in all three plies. The temperature drop through each of the glass plies is half what it would be if the heat entering the ply were all leaving the plate instead of being entirely absorbed.

The time required by the windshield to approximate the condition of steady state of rise of temperature was determined graphically to be about 18.5 minutes. Temperatures corresponding to curves *A*, *B*, and *C* were plotted against time, using more points in the panel than surface and interface points. These were faired into straight lines determined by curve *D* and by a slope equal to the quotient of heat input by thermal capacity of the panel.

The curves of Fig. 2 (except *D* and *E*) scale vertically to whatever value of *Q* is supplied to the windshield. No temperature gradient in the forward glass during ground warm-up is as high as the initial gradient which is 71 per cent of the glass gradient in steady-state flight *E* with all of the heat input going forward due to negligible heat interchange with cockpit air.

During the 6-year interval since this partly graphical summation process was developed, it has proved useful in a surprising number of ways. For example, the number of windshield failures in service has been considerably smaller than it would have been without the guidance of this solution. This is especially true as to bubbling of vinyl plastic because of overheat at the hot spots of the panel surface. Guesswork is considerably reduced for changes of the temperature control system for better operation.

Flat-Plate Case

The Heaviside type of solution for an insulated plate of thickness x_1, electrically heated on one surface, as developed in transmission line terms by engineers of the Boeing Airplane Company, may be written in the form

$$t_h=(Qx_1/\pi^2k)f_h(z)=(Qx_1/\pi^2k)\left[z+\pi^2/3-2\left(e^{-z}+(1/4)e^{-4z}+(1/9)e^{-9z}+\ldots\right)\right] \quad (14)$$

$$t_c=(Qx_1/\pi^2k)f_c(z)=(Qx_1/\pi^2k)\left[z-\pi^2/6+2\left(e^{-z}-(1/4)e^{-4z}+(1/9)e^{-9z}\ldots\right)\right] \quad (15)$$

In this form, $z=\pi^2\alpha\theta/x_1^2$ (reciprocal to *y* for the infinite medium case except as to numerical constant); t_h and t_c are hot and cold surface temperature histories; $f_h(z)$ and $f_c(z)$ are dimensionless functions of time.

As seen in curves $f_h(z)$ and $f_c(z)$ of Fig. 3, the rapid convergence of both of the exponential series places $f_h(z)$ and $f_c(z)$ at practically the same distance on the $f(z)$ scale respectively below and above their asymptotes, when $z=\pi^2/3=3.29$. Both are vertically below or above their asymptotes to the extent of 4.53 per cent of the $\pi^2/6$ vertical axis intercept of the lower asymptote. The hot-surface function is substantially a parabola $(4\pi z)^{1/2}$ to this point (temperature rise proportional to the square root of time) because the hot side is just discovering that the medium is not semi-infinite. In the practical case of a 3/16-inch electrically heated glass under test for power constants this time is about 16.5 seconds. It would be four times as long for a 3/8-inch glass. But the intercepts of the asymptotes with the vertical axis and the temperature rise at each surface would all be twice as great, while the slope of the asymptotes would be half as great for the same heat input rate.

In contrast with the 16.5-second time and the 2-to-1 ratio of the asymptote intercepts that hold for a 3/16-inch unlaminated unilaterally heated forward glass ply, the time to steady state of rise for the internally heated windshield panel is about 65 times as long and the ratio of asymptote intercepts of heated internal coating and panel aft surface is about 1.25. The forward ply is able to maintain infinite line behavior for only about 1 minute before it begins to follow a sweeping temperature curve toward its asymptote.

The particular value $z=\pi^2/3$ may be viewed as a time constant for any kind of insulated flat plate receiving suddenly started constant heat input at one surface. It marks the time at which conversion has practically been completed from a transient state having a complex internal temperature distribution to a state of steady rise in temperature with the temperatures within the plate fitting a parabola having its vertex on the cold side. In the early stages, temperature history at any given internal point not too near the cold surface has an S-curve shape. Curves for all values of *x* but zero start horizontally.

Timing Method for Power Constant Determination

Prior to the time at which information became available on the Heaviside analysis of a flat plate that forms the basis for industrial measurement of the thermal quality of glass after application of an electrically conductive coating, equation 8 supplied the basis for a different method of evaluation. The location of hot spots on glasses had for some time been determined before lamination by applying a thin coating of a white waxlike material to the conductively coated side. Application of full power identified hot spots by observation of first melting of wax but provided little information on the magnitude of the heat production at hot spot or other points.

Temperature rise proportional to the product of local heat input by the square root of time could be used to compare local heat inputs by determining the time required to rise from a uniform starting temperature to wax-melting point. It would not be practicable to reduce working time to the $16^1/_2$ second during which the 3/16-inch glass acts like an infinite medium on its heated side. But the reflection laws equations 12 and 13 show that no reflection occurs when $m=1$. A long glass line can be established by laying the glass to be tested, coating side up, on a stack of similar glasses with thermal contacts established by water to which a little ethylene glycol has been added to reduce its $kc\rho$ to that of glass.

Single Heat Pulse into Insulated Flat Plate

If a boot-wiring configuration were not too radically different from a plane heat source and if, as previously suggested, reversal of initially aftward heat flow were greatly limited by suitable design, it might be represented by, say, a 3-second plate, one that approximates steady rise of temperature after three seconds. The effect of three equal quantities of heat put into one surface at quite different rates is shown in Fig. 1. Heat input rate of *A* curves is 2.5 times that for *B* curves, for which the rate is 2.5 times that for *C* curves.

Stopping a heat flow is equivalent to the fiction of permitting it to continue while superposing on it from the moment of actual stoppage the temperature decreasing effects of an equal sink created at that moment at source position.

The dimensionless curves of Fig. 3 are converted to temperature curves and their associated asymptotes by equations 14 and 15. The plate will level in temperature about 3 seconds after the end of the pulse, slightly longer for the highest input rate, curves *Ah* and *Ac*, the respective hot and cold surface histories for the fastest pulse. The level temperatures will be the same for pulses of equal heat content.

The graphs of Fig. 1 show that the plate temperature levels for the fastest pulse about $5^1/_2$ seconds after pulse start; the slowest about $12^1/_2$ seconds. It may be surprising that the fast pulse, 6.25 times as strong as the slow one and of correspondingly short duration, causes a little less than twice as great a rise in hot surface temperature. The possible trend of peak temperature for other pulses of equal heat content is indicated by sections of a curve through the peaks of the curves of source temperature.

Curves of deicer boots having widely spaced wires as the heat source and, having other interferences with heat transmission, would be expected in a thermal open-circuit test to give a weather surface temperature curve somewhat like the *Cc* curve of the simple plate and a source temperature curve somewhat like the *Bh* curve extended in time. The 0.020-inch breeze ply of such a boot ought to reach steady state of rise in less than 1 second on the basis of a time constant computed from $z=\pi^2/3$. Indications from limited data are that it behaves timewise more like a 0.050-inch rubber plate that has a 5-second time constant.

Attention is directed toward the lack of sudden changes in the cold-side temperature curves of Fig. 1 in contrast with marked changes in slope of the curves of source temperature.

Effects of Cooling

Until analysis of a plate subjected to forced convective cooling is available, discussion is limited to the following:

1. In a low-heat transfer coefficient situation, a plate such as that of Fig. 1 would have lost about twice as much heat at 3 seconds after the end of the slow pulse as at 3 seconds after the fast pulse, merely because it had been cooled for a longer time. The time average of cold surface temperature until leveling time is about the same for all three cases, giving substantially identical heat losses per unit time when averaged over the respective intervals.
2. For high-heat transfer, the heat loss rate might, as the temperature rises during slow input, reach heat input rate. The resulting constancy of temperature during the rest of the pulse might be at too low a level. But the fast arrival of heat at the surface, as shown by the greater slope of the *Ac* curve of Fig. 1, gives the fast pulse a greatly improved chance of producing useful temperature. There are also indications that

ice may be released without melting, because of high thermal stresses in brittle ice when heat can be supplied rapidly to its contact surface so that the ice is stressed by its own high-temperature gradient.

Time Constant in Grounded Case

Besides the time constant defined as time to reach steady state of temperature rise for an insulated plate, or the nearly equivalent time to level temperature after the pulse, there may be some value in another constant. The time required to reach steady state for a grounded transmission line or the time required after pulse end for the plate to level to the locked down temperature of its forward surface determines the shortest time within which all the heat pulse can be transmitted by this plate to the forward surface. Measurement of either of these time intervals would require a heater element that has a substantial increase of resistance with temperature.

References

1. Intermittent Heating for Aircraft Ice Prevention, **Myron Tribus.** *Transactions,* American Society of Mechanical Engineers, New York, N. Y., vol. 73, Nov. 1951, p. 1117.

2. Standard Handbook for Electrical Engineers, **A. E. Knowlton.** McGraw-Hill Book Company, New York, N. Y., 1949, seventh edition, section 13–6.

3. Constant Current D-C Transmission, **C. H. Willis, B. D. Bedford, F. R. Elder.** *Electrical Engineering,* vol. 55, Jan. 1935, p. 102–08; discussion Aug. 1935, p. 882–83.

4. Heat Transfer (book), **Max Jakob.** John Wiley and Sons, Inc., New York, N. Y., 1949, vol. 1, chap. 13.

5. Handbook of Engineering Fundamentals, **O. W. Eshback.** John Wiley and Sons, Inc., New York, N. Y., 1936, section 2, p. 124.

◆

No Discussion

Problems Relating to Interconnections of Large Pulp and Paper Mills with Large Utility Power Systems

H. A. ROSE
MEMBER AIEE

H. E. SPRINGER
MEMBER AIEE

THE requirements for electric power in modern pulp and paper mills are increasing at greater rates than the requirements for thermal power. In consequence, the electric power that can be generated by the normal flow rate of process steam is becoming inadequate. Low-cost hogged fuel and mill waste are also no longer available in abundant quantities to generate the increasing unbalance in electric load with condensing steam. Therefore, mill operators are finding it more economical to purchase a portion of their electric energy from utility power systems rather than to operate their own generating equipment on more expensive fuels and with consequent heat loss to the condensing water.

This requires interconnecting the principal power busses in the mill, with circuits of the utility for joint or parallel operation, for which Figs. 1 and 2 represent typical arrangements. The power tie often requires major modifications in the mill electric system. New problems and methods of plant operation are usually encountered.

The total cost to the mill to make a major utility connection is usually considerable; final cost estimates are often much higher than anticipated during initial stages of investigation. Appreciation of the more important factors of installation and operation of such power ties is therefore desirable.

The authors' primary interests are directed at benefiting the mill, namely:

1. The interconnecting power circuitry shall be adequate to meet the rigid requirements for continuity of service to the mill's principal product producing departments.
2. The interconnection shall be made at minimum initial cost commensurate with the mill's needs for prime circuit flexibility and future growth requirements.
3. Such changes as may be required in the mill's internal power-distribution system shall be made at minimum cost commensurate with the requirements for continuity of service and safety to equipment and personnel.
4. The flow of electric energy to the mill may be controlled to obtain minimum cost rates under terms of the power contract.
5. Fault conditions within the utility circuits shall be relieved or disconnections made so that prime mill processes are not adversely affected.
6. The resulting system shall operate economically efficient in all its principal parts.

Relationship of Electric and Thermal Power

The principal reason that mills purchase power is one of economics, i.e., a portion of their electric energy may be purchased for less cost than would be required to generate it.

The availability of low-cost power of adequate capacity and reliability has had profound influence on the quantities for which mills have considered it safe and economical to contract. To be of most value, purchased power must be of good reliability and have good voltage regulation at the delivery point to the mill. The power system should be capable of paralleling the normal growth of the mill power needs.

The mills require large amounts of process heat energy in the form of low-pressure steam. In the United States this energy is generated by fuel-fired boilers. It is unlikely that in the foreseeable future this energy can be as economically purchased as electric energy even though highly efficient electric boilers are available for heat conversion. It should be recognized that utility electricity, except in the form of heat, cannot serve the thermal processes of the mill. Heat energy must be generated by burning mill waste or purchased fuel of some kind. However, as a utility product, electricity can fulfill the energy requirements for all the mills' wheel-turning and mechanical power-consuming machinery.

It is important to recognize that most pulp and paper mills do have fairly large-size electric generating plants and find it economical to operate them. This condition results from the fact that the incremental investment and operating costs required to add electric power generation to the already required steam boiler plant

---

Paper 54-243, recommended by the AIEE Power Generation and Industrial Power Systems Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 17, 1954; made available for printing May 8, 1954.

H. A. Rose is with Westinghouse Electric Corporation, Seattle, Wash., and H. E. Springer is with Rayonier, Inc., Port Angeles, Wash.

The authors wish to express their appreciation to K. H. Schiffleger and G. E. Grosser for valued suggestions and assistance given during the preparation of this paper.

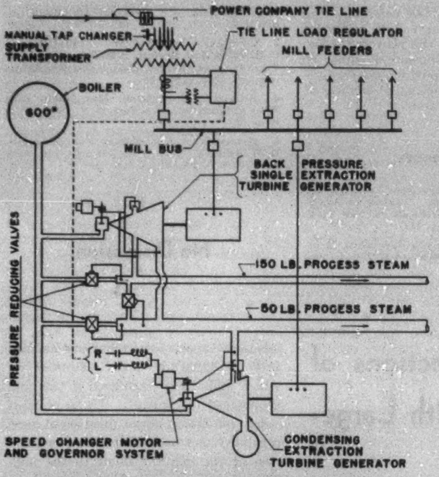

POWER COMPANY TIE LINE
MANUAL TAP CHANGER
SUPPLY TRANSFORMER
TIE LINE LOAD REGULATOR
MILL FEEDERS
BOILER
600*
MILL BUS
BACK PRESSURE SINGLE EXTRACTION TURBINE GENERATOR
150 LB. PROCESS STEAM
50 LB. PROCESS STEAM
PRESSURE REDUCING VALVES
R
L
SPEED CHANGER MOTOR AND GOVERNOR SYSTEM
CONDENSING EXTRACTION TURBINE GENERATOR

Fig. 1. Typical steam and electric power apparatus diagram for a pulp mill, showing a utility power system interconnection

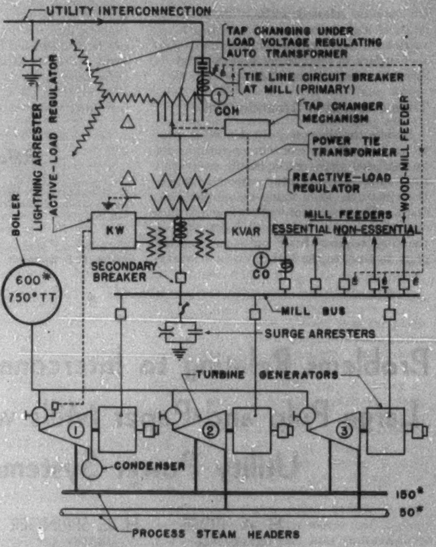

UTILITY INTERCONNECTION
TAP CHANGING UNDER LOAD VOLTAGE REGULATING AUTO TRANSFORMER
TIE LINE CIRCUIT BREAKER AT MILL (PRIMARY)
COH
TAP CHANGER MECHANISM
POWER TIE TRANSFORMER
REACTIVE-LOAD REGULATOR
MILL FEEDERS
ESSENTIAL
NON-ESSENTIAL
WOOD-MILL FEEDER
LIGHTNING ARRESTER
ACTIVE-LOAD REGULATOR
BOILER
KW
KVAR
SECONDARY BREAKER
CO
600*
750°TT
MILL BUS
SURGE ARRESTERS
TURBINE GENERATORS
①
②
③
CONDENSER
150*
50*
PROCESS STEAM HEADERS

Fig. 2. Typical steam and electric power diagram for a pulp mill, including: utility tie with tap-changing-under-load transformer equipment, regulators for control of kw and kilovar in the tie, relay means (COH) for high-speed severance of the tie with simultaneous tripping (T) of nonessential mill feeders, and lightning and surge protective equipment

are small enough to make it economical to generate the full amount of electric energy that mill waste and low-cost hogged fuel will produce over that required for the thermal processes.

This is accomplished by boilers and generating equipment designed to operate in conjunction with the mill's processes but at elevated temperatures and pressures, augmented by steam superheating and condensing equipment. In this way the normal flow rate of process steam generates a large portion of the mill's electric power requirements. If more electric power than this is required to turn the mill, the additional amount must be generated by steam flow to the condenser. If this increment has been generated by waste fuel, a satisfactory heat balance can be assumed, or conceivably the mill might have an excess of electric power for sale. If there is a deficiency of waste fuel to balance electric requirements, fuel oil, or other types of fuel, will have to be burned in proportionate amounts and at comparatively high cost. It should be recognized that heat liberated at the condenser is lost.

It is at this point in the balance of economic factors involved that the utility may be able to offer sufficiently attractive rates for specific amounts or blocks of electric energy. That this balance point has been reached is evidenced by the large number of pulp and paper mills which purchase all or a substantial portion of their electric energy.

Power-System Layout and Switchgear Requirements

Many papers have been written on electric power systems for industrial plants.[1] It is not in the scope of this paper to discuss types of systems specifically applicable to pulp and paper mills. The power system for a properly laid-out mill will most generally use simple radial feeders, simple loop feeders, or a combination of both, feeding from the principal sources of supply, to step down load-center transformer stations near centers of load in different departments of the mill.

A utility tie usually will not require changes in the distribution system if it is of modern load-center design. If not of such design it is likely that major changes will be required to insure process power reliability and safety to equipment and personnel.

A strong power tie is very apt to require major modifications in engine-room power circuitry and switchgear. These may entail replacement of old circuit breakers with new equipment having adequate short-circuit interrupting ratings to meet the increased duty imposed.

It is important to understand why costly plant modifications may be required so that management can be properly informed and adequate funds provided. A few principal factors are usually responsible:

1. The addition of a new source of power in parallel with existing generators may increase the internal short-circuit capacity of the mill's circuits so that higher fault-capacity circuit breakers are required, even though the tie may involve no increase in mill load.

2. The mill's switchgear may have inadequate fault capacity for the existing plant if additional generating capacity has been added without proper corrective measures being taken at the time.

3. An isolated turbine-generator plant, or one having a weak power tie, when subject to short circuit will rapidly diminish its fault-current output. By the time existing slow-speed circuit breakers open their contacts the fault-interrupting requirement of the system has decreased to a comparatively low value, thus minimizing the duty imposed. When such a plant is connected directly with a strong power tie, the tie does not diminish its fault-current output with time. In such cases, existing switchgear may be entirely inadequate for the

new conditions. Fig. 3 is a typical illustration of the increased short-circuit interrupting requirements produced by a strong utility tie.

4. Other deficiencies of existing switchgear may be:

   a. Inadequate momentary current-carrying capability under fault conditions.
   b. Inadequate current-carrying capacity for normal loads.
   c. Old obsolete design, manually operated, unreliable.

It should be emphasized that there is nothing fundamentally wrong in taking advantage of the inherent fault decrement characteristics of a-c generators to make continued use of old slow-speed but otherwise satisfactory switchgear, when this can be properly done. It is this characteristic, coupled with intentional time-delayed relay settings, that has been responsible for years of very satisfactory and safe performance of otherwise inadequate circuit-breaker equipment in many of our older mills. When the occasion justifies, the authors will take advantage of the decrement characteristic and justify continued use of existing good quality switchgear which otherwise might be replaced at considerable unnecessary cost.

A detailed engineering analysis of the mill's power system should be made to determine the proper equipment and layout of circuits required for any power tie. Such studies produce best results when made in close co-operation with the mill's engineering personnel.

Capacity of the Power Tie

The power tie will usually be made through a supply transformer, for stepping the utility voltage down to one of the mill's principal bus voltages. For mills having more than one voltage (other than 480 volts) the tie will usually be made to the bus of highest voltage.

Normal tendency is to size the transformer to meet the immediate contract load. However, it is well to consider near future power requirements and size the transformer and primary circuit breaker accordingly. This point is of particular importance if the mill is to provide the connecting equipment. Otherwise, installation of undersize equipment will later result in overly complicated and expensive switch structure additions. The power supply station for a mill should always be of simplest design keeping to a minimum the number of insulator strings, bus supports, disconnecting switches, and other devices that can collect deposits from mill fume. The best way to avoid trouble with such devices is to eliminate them by design.

Regardless of the tie capacity that is decided on, serious consideration should be given to the future probable transformer bank rating. Based on this rating, the secondary power circuit arrangement and fault-interrupting capacity for the mill's main bus switchgear should be determined. Attention to this important matter will obviate the necessity and embarrassment of having to make expensive corrections at some early future date.

Operation of Mill's Turbine Generators Prior to a Utility Tie

The isolated mill in addition to producing its process steam and electric power must also maintain its own 60-cycle plant frequency. A typical mill might have two generating units; see Fig. 1. One could be a single-extraction, back-pressure unit, the other a single-extraction, condensing unit. Since process steam flow rates are quite variable, this steam in passing through the turbines generates proportionately variable power. Because electric load requirements are also variable and bear no fixed relation to process steam demands, there will be times when the total kilowatts (kw) of electric load will exceed process steam capability. At other times process steam capability may exceed electric load requirements. It is obvious that continuous ideally balanced operation is not attainable.

To provide for this disproportionate relationship, mill turbines are equipped with elaborate steam-flow and speed-governor controls.[2-4] At least one unit will have a steam condenser. Automatic pressure-reducing steam-flow valves will be installed between boiler pressure and the process lines to supplement turbine flows as illustrated by Figs. 1, 2, and 4.

The condensing unit, Fig. 1, would be used to establish mill frequency, it being set on speed-governor control to hold 60 cycles. The other unit would run in synchronism as a nonspeed regulated machine generating power in proportion to process steam flow through it. Both turbines would be arranged with extraction and exhaust process-pressure regulators, adjusted for parallel steam-flow operation according to process needs.

When process steam requirements are in equivalent balance with electric load, the extraction and exhaust pressure controls will adjust steam-flow rates through sections of turbine blading so that proper process line pressures are maintained. In this case, there will be no flow through the reducing valves and there will be no steam, except for a small amount of cooling steam, passing to the condenser.

Should electric load demand exceed equivalent process steam requirements, the frequency will lower slightly. This affects the speed governor of the condensing unit which actuates the turbine's control valves, admitting additional steam to re-establish frequency and to generate

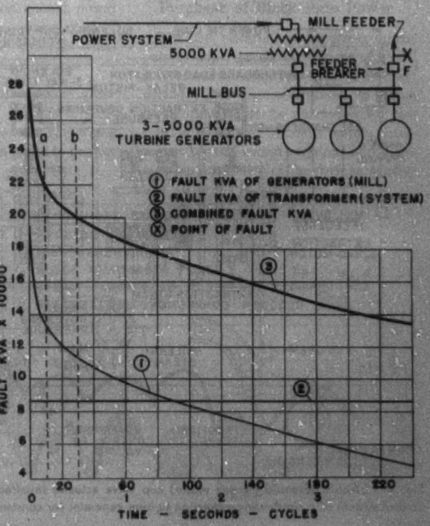


Fig. 3. Increased fault kilovolt-amperes on mill's engine-room power circuits ① before and ② after installation of a utility power system tie

a—High-speed tripping of modern circuit breaker at F
b—Delayed tripping of slow-speed circuit breaker at F

the increased load. The increased flow tends to raise process pressures above normal, resulting in essentially instantaneous automatic readjustment of both turbines' extraction controls so that the excess steam required to generate the additional electric load passes to the condenser. However the extraction pressure remains constant because the extraction servomotor will also operate to compensate for the increased steam flow by passing it on to the condenser.

Should electric load demand be less than equivalent process steam requirements, the frequency will raise slightly, affecting the speed governor and associated control valves to reduce steam flow. The decreased flow lowers process line pressures which opens the pressure-reducing valves to make up the deficiency direct from boiler pressure.

Different types, combinations, and numbers of turbine-generator units will be used to meet different mill's processes and power requirements. The underlying fundamentals of operation will be essentially as described.

Operation of Mill's Turbine Generators with a Power System Tie

The generators of the nonisolated mill must run in synchronism with the utility. Since the generating capacity of the power system is usually large compared to mill capacity and is regulated at essentially constant frequency, the mill's turbine speed governors have no control over system frequency. Their frequency-controlling ability is lost at the instant of synchronization.

If, at the instant of synchronization, both systems are running at exactly 60 cycles, there will be no interchange of power, but the turbine governors will instantly lose all ability to control speed; in fact, there is normally no system speed change to affect them. If a system speed (frequency) change does occur, the turbine governors will be affected and any units that are set for speed regulation in this frequency range will experience decided changes in steam flow and load.

If the system frequency raises slightly, the governors on speed control will reduce steam flow, unloading the units. A reduction in system frequency will increase load on the same units. The effect is very pronounced. For example, a 1-per-cent system frequency change might change load on any of the mill's speed-regulated units about 25 per cent, but the utility would experience only about 1-per-cent load change.

Assume now that there is no power interchange at the instant of synchronization. If the mill load increases the utility will absorb the entire increment since there will be essentially no frequency change and, consequently, no governor action to increase steam flow. Fluctuations in mill load will result in equal fluctuations in utility load but the output of the mill's generators will remain unchanged. A decrease in mill load from the initial value assumed (zero interchange) results in a reversal of power in the tie so that the mill generators help carry the external system load.

It is common to say that the mill generators carry a portion of mill load or some additional portion of the utility load. Actually, they assume a portion of the total system load in proportion to their torque energies.

If the utility maintains normal frequency it can exercise no control over the power taken by the mill. The magnitude and direction of power flow in the tie can be determined and regulated only by the prime movers of the mill's own generators. To change the load generated by these units requires:

1. A change in process steam flow as initiated by extraction and back-pressure controls. This applies primarily to extraction and back-pressure machines not on speed-governor control. Since process steam-flow rates are governed by process requirements, noncondensing extraction back-pressure machines are not often used for load regulating purposes.

2. A change in adjustment of the speed-changer mechanism, i.e., speed-governor setting, so that for the same system speed more steam is admitted by the governor control valves. This applies primarily to condensing machines or other machines operating on speed-governor control.

Control of Real Power Flow in the Tie

The magnitude and direction of real power flow in the tie will usually be regulated by a condensing turbine generator operated on either manual or automatic load control. Automatic load control is obtained by a kw load regulating device connected to current and potential transformers in the tie circuit. This device actuates appropriate secondary controls to alter the turbine speed-governor settings in response to load changes in the tie. This action varies the steam-flow rate to the turbine to meet the changes in load requirements. Operation of the system is only secondarily affected by extraction conditions, since steam not required for process purposes flows to the condenser. The turbine blade paths and the condenser area must necessarily have capacity to handle the quantities of steam required for both process and load-regulating purposes. Such load-regulating systems can be used to hold purchased power constant or to limit it to a predetermined value.

Two principles may be employed to alter the speed-governor setting for automatic load control purposes:

1. The load-regulating device may operate the usual synchronizing or speed-changer

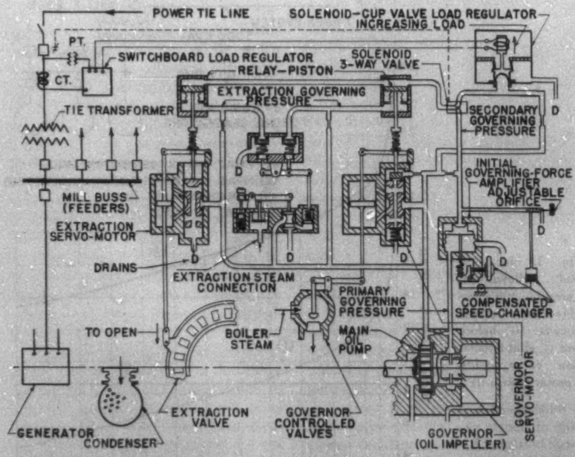

POWER TIE LINE
SOLENOID-CUP VALVE LOAD REGULATOR
INCREASING LOAD
PT.
SWITCHBOARD LOAD REGULATOR
SOLENOID 3-WAY VALVE
RELAY-PISTON
CT.
EXTRACTION GOVERNING PRESSURE
TIE TRANSFORMER
SECONDARY GOVERNING PRESSURE
D
INITIAL GOVERNING-FORCE AMPLIFIER ADJUSTABLE ORIFICE
MILL BUSS (FEEDERS)
EXTRACTION SERVO-MOTOR
DRAINS
EXTRACTION STEAM CONNECTION
PRIMARY GOVERNING PRESSURE
COMPENSATED SPEED-CHANGER
TO OPEN
BOILER STEAM
MAIN OIL PUMP
GOVERNOR SERVO-MOTOR
EXTRACTION VALVE
GOVERNOR CONTROLLED VALVES
GOVERNOR (OIL IMPELLER)
GENERATOR
CONDENSER

Fig. 4. Electric solenoid (torque motor) cup valve actuator installed in the hydraulic speed control system of a condensing-extraction turbine generator for constant kw tie-line load control

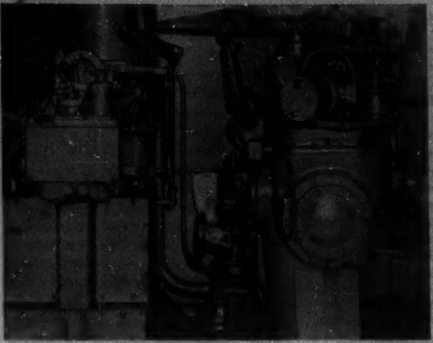


Fig. 5 (left). High-speed, heavy-duty speed-changer motor and governor-head mechanism installed on a condensing-extraction turbine generator for constant kw tie-line load control

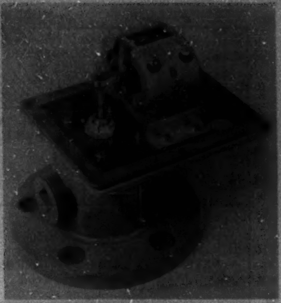


Fig. 6. Electric solenoid cup valve load regulator of Fig. 4

motor to increase or decrease governor spring tension, thereby altering the steam flow and power output of the unit. See Fig. 5.

2. The load-regulating device may operate an electric torque motor or solenoid to change the net effective speed-governing oil pressure, thereby altering the steam flow and power output of the unit. See Fig. 4.

The first system is particularly applicable to interconnections where load changes are gradual and of moderate magnitude. If load changes are more violent and recurring, a high-speed heavy-duty speed changer motor and governor-head mechanism will be required to obtain sufficiently fast operation and satisfactory apparatus life. This system is readily adaptable to old or existing turbines where moderate accuracy of load control permits its use. Provision must be made to prevent the speed-changer motor from closing the steam valves in the event of a break in the power tie. Fig. 5 shows a high-speed rotating regulator applied to the governing system of a condensing-extraction turbine generator for constant kw tie-line control to a pulp mill.

The second system is particularly applicable where load swings are fast and of considerable magnitude, or where a high degree of control accuracy is required. This system has the advantage of not disturbing the normal settings and functioning of the turbine's usual speed governor. In the event of a break in the power tie, the governor starts controlling at its normal preset full-load 60-cycle frequency.

In operation the load-controlling solenoid, shown in Figs. 4 and 6, applies force to a governor oil-pressure transformer cup valve. Very slight movements of this cup valve varies the governor's secondary control pressure which acts on a servomotor piston to adjust the admission valves that regulate the steam input to the turbine. Since this device acts directly in the turbines' hydraulic speed control system, the response is essentially instantaneous and consequently it maintains accurate control of the tie-line load.

To hold constant load (block purchased power) in the tie it is necessary that steam pass to the condenser during certain intervals. This condition prevails when:

$$\text{Process kw capability} + \text{purchased kw} < \text{total kw of mill load}$$

In determining the block power rate it is important to evaluate the net condenser heat loss during these intervals and balance it against the sale value of hogged fuel. This fuel if not burned or sold will accumulate. It will usually be advisable to contract for just enough power to keep waste fuels from accumulating.

Fig. 7 shows the time intervals when heat is lost at the condenser because

$$\text{electric load requirements} > \text{process capability} + \text{block purchased power}$$

The intervals are also shown during which steam flows directly from boiler to process pressures, because

$$\text{electric load requirements} < \text{process capability} + \text{block purchased power}$$

The curves are not necessarily typical of continuous mill operation. They were assumed to illustrate the different conditions incurred when an approximate maximum value of block power E_p is purchased.

Mills producing ground wood pulp may regulate control of purchased power by a tie-line regulator arranged to vary grinding stone pressures. Excessive short-time load demands exceeding the correcting capability of the regulating system may be handled by stopping a grinder automatically or by opening nonessential load feeders.

Effect of Woodmill Operation on Purchase of Block Rate Power

The base electric load for powering the pulp or paper producing portion of a mill is reasonably constant over a 24-hour period. The kw magnitude of block power that might be purchased for this portion of the mill's load is readily determined by the excess required above extraction capability of the turbine genera-

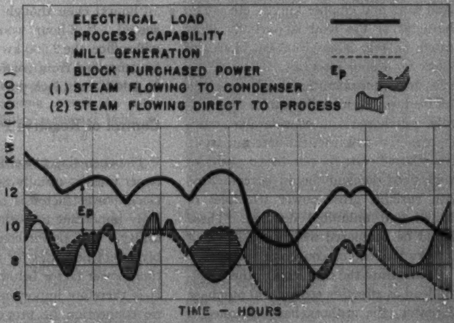


Fig. 7. Power diagram for a pulp mill showing intervals of condensing power generation and intervals of boiler steam flowing direct to process when block purchased power E_p is at a near maximum economic valve

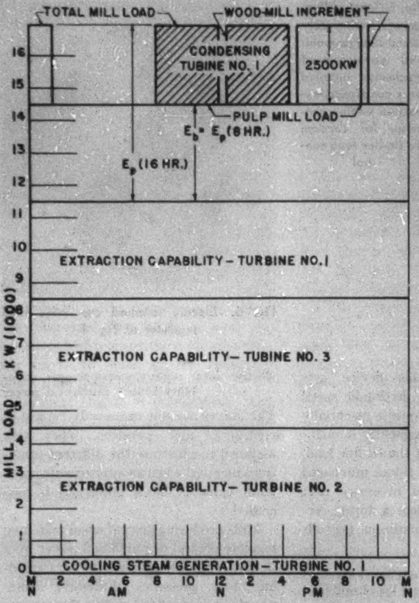


Fig. 8. Power diagram for a pulp mill showing average load division between generators. 14,500-kw average pulp mill load. $E_b = 3,000$ kw of block purchased power for pulp mill operation. 2,500-kw woodmill load for 8-hour and 16-hour operation. Based on power data of Fig. 9. Most economical operation with the woodmill loads is obtained when:

8-hour woodmill operation = carry woodmill load on no. 1 turbine. Condensing and purchase no additional power above E_p

16-hour woodmill operation = purchase additional block power totaling E_p (16 hours)

tors with allowance made for generation from necessary cooling steam. The economic feasibility depends primarily on the mill's cost rate to produce block power from condensing steam compared to that for which it may be purchased. Fig. 8 shows this block of power for a 24-hour period as E_b.

The load of the wood preparation plant, however, is quite variable because of the nature of the work done and the types of machines required to reduce the logs to suitable form for process purposes. The woodmill is also usually operated one or two 8-hour shifts per day with time out for lunch and relief periods, for changing band saws and chipper knives, etc. The character of the load is partially illustrated in Fig. 8 as 2,500-kw blocks of power for 4-hour intervals.

Separate consideration must be given this load to determine if it can be economically purchased; a considerable number of factors will enter the determination. The relative costs for purchased power of block fixed demand should be compared with variable gross power schedules of limited and unlimited demand. These should be compared with the mill's cost to carry the load on condensing steam.

To illustrate the character of the woodmill load as affecting economical purchase of power, the authors have assumed a typical case. The power cost data, load values, and days of operation are given in Figs. 8 and 9. Both 8-hour and 16-hour woodmill operation are considered. Fig. 9 shows the incremental woodmill cost for purchased power plus condensing power, as the mill's total block purchased power is increased from the 14.5-kw ($E_b = 3,000$ kw) to the 17.0-kw level.

It is important to note that 8-hour woodmill operation on condensing power costs considerably less than 16-hour operation on block purchased power. Also, if 16-hour operation is required, it is more economical to purchase the 2,500-kw woodmill increment continuously (E_p, 16 hours) even though it is not all consumed. If 8-hour woodmill production will suffice, the 2,500-kw increment should be generated from condensing steam and no additional block power purchased.

Control of Reactive Power in the Tie

The operating power factor of pulp and paper mills is usually 80 to 95 per cent. Good power factor is becoming increasingly important to both the mill and the utility. The power contract will probably inflict penalties for average or short-time power factor below specified values. It is therefore important that the power tie connection be properly co-ordinated with the mill's circuits and facilities so that adequate control of power factor can be assured.

When two independent sources of generation are to be connected, differences in normal operating voltages at the tie point are likely to exist. It is essential that the tie transformer have suitable range of voltage taps, Fig. 1, above and below the normal value so that the connection can be made at essentially zero voltage difference. The transformer should be provided with a no-load type of tap changer for manual operation from outside the transformer case.

If this is not done the system with highest voltage will deliver reactive power to the other, thus altering the normal (unity) power-factor state of the tie. The magnitude of this reactive interchange will be proportional to the voltage difference. The condition of unity tie power factor means that the mill is carrying its own reactive load as is the utility.

The power contract may provide for the utility to carry a portion of the mill's reactive load. To accomplish this it becomes necessary to raise the utility's input voltage through the tie transformer, or to lower the mill's bus voltage a like amount.

It is very undesirable for the mill to operate at subnormal voltage. Normal bus voltages insure the most reliable operation of the large amount of electric apparatus; most efficient operation of the many motors is obtained. Maximum mechanical output and product producing capability are also assured. Low-voltage conditions are often associated with poor power factor or with poor facilities for proper control of reactive power.

If the tie transformer has been provided with an adequate range of voltage control taps, it will be possible to choose one that will produce the proper directional flow of reactive power and still allow the mill generators to operate at essentially normal bus voltage. Slight changes in generator voltage can then be made to keep the tie power factor within prescribed limits.

For best operation, mill generators should have their voltages controlled by individual modern voltage regulators with cross-current compensation to insure proper division of reactive load between machines and to improve their load-carrying capabilities.

A power system which delivers low voltage at the tie point during its heavy load periods sheds its reactive to the mill generators, reducing their real load capacity. To alleviate this situation without lowering generator bus voltage it becomes necessary for the mill to either:

1. Carry the additional reactive load through the peak with its normal load generators.

2. Place additional generators on the bus to alleviate the overloaded condition.

3. Split the mill generating circuits, placing the tie on non essential mill load and operating isolated on the essential portion.

4. Disconnect the tie long enough to change transformer taps.

Similarly, a system which delivers high voltage can absorb more than a normal amount of the mill's reactive load and becomes responsible for the mill incurring low power factor or reactive demand penalties.

The well-regulated heavier power system circuits usually have sufficiently well-maintained voltages to permit setting the tie transformer on a tap between the high- and low-voltage variations. For such systems, only slight changes from normal generator voltage are required for adequate reactive load control.

For such systems, tie-line reactive load regulators or load limit controllers may be used in conjunction with the generator voltage regulators to hold the reactive power automatically within prescribed limits. The controls should be arranged to become inoperative when the bus voltage reaches predetermined values. Otherwise, abnormal load conditions may result in excessively high or low mill voltage, with the regulators attempting to hold the tie-line reactive load within the prescribed limits.

Where interconnections are made to weak or poorly regulated power circuits or to major systems subject to wide voltage fluctuations, it may be necessary or desirable to supplement the tie transformer with tap-changing-under-load equipment. See Fig. 2. With this equipment the generator busses may be operated at normal voltage and the tie line allowed to vary in conformance to the operating characteristics of the system. The tap changer position may be automatically controlled to regulate the flow of reactive power in the tie or to limit it to zero value.

A mill having limited fuel supply but ample reactive (kilovar) generating capability might advantageously purchase real (kw) power and simultaneously assist the utility by supplying it reactive power for improving system voltage and stability. The fact that the directions of real and reactive power flows are in a sense reversed to each other in the tie does not prevent satisfactory and automatic operation with this arrangement.

Regulators for Controlling Real and Reactive Load in the Tie Circuit

The real and reactive load regulators usually form a section of the mill's main switchboard. The regulators' load-sensing elements are energized by current and potential transformers of the tie circuit. The sensing networks control the output of the regulators which in turn cause operation of the:

a. Turbine's speed changer motor for kw corrections, or

b. Turbine's hydraulic cup valve actuator for kw corrections.

c. The generator voltage regulator compensators for var corrections, or

d. The transformer underload tap changer for var corrections, in the power tie circuit.

Load regulators may be of the following types which characterize their principles of operation:

1. Magnetic amplifier controlled power output.[5-7]

2. Rotating regulator controlled power output.[8,9]

3. Electronic tube controlled power output.[9]

4. Balanced beam relay controller; see Fig. 10.

5. Balanced bridge circuit controller.

Of these, types 1, 2, and 3 have the fastest speeds of response, and are capable of meeting the accuracy requirements of most tie circuit applications. Their circuits are usually somewhat complicated, which is reflected in higher equipment cost compared to types 4 and 5. They are preferably used to control systems a, b, and c. Types 4 and 5 are simple, of lower cost, and sufficiently accurate for most applications. They are best used to control systems a and d.

When energy is purchased on a block kw demand basis, minimum cost occurs when, at the end of each demand period, the number of kw hours permitted by the

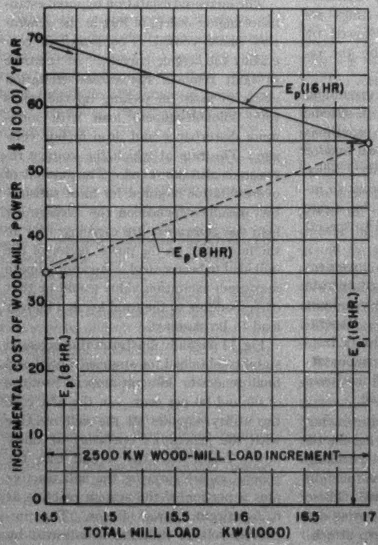


Fig. 9 (left). Variation in cost of purchased power+cost of condensing power to carry 2,500-kw woodmill loads of Fig. 8

Block purchased power=2.5 mils per kilowatt-hour, basis 360 days per year

Condensing power=5.0 mils per kilowatt-hour, basis 350 days operation per year

E_p (16 hours)=5,500 kw (Fig. 8—noncondensing operation)

E_p (8 hours)=3,000 kw (Fig. 8—condensing operation)

Fig. 10. Balanced beam type of kw or kilovar load regulating relay. The relay energizes auxiliary speed-reversing control contactors which energize the motor of the turbine speed changer for kw-load control and the tie transformer tap-changing-under-load mechanism for kilovar load control

demand interval has been consumed. A high-accuracy, fast-response, constant kw regulator approaches the needed requirements. However, it must be set to regulate at slightly less than the permissible rate to avoid the possibility of exceeding demand.

One solution to this problem is to superimpose the control of a kw-hour integrator on that of the constant kw regulator. The integrator continuously recalibrates the regulator according to the actual as compared to the permissible rate of energy consumption, so that at the end of each demand period essentially the total allowable kw-hours have been purchased.

The need for reactive power regulators has not been as great as for the real power types. They can be provided when required and operate on essentially the same principles as real power regulators.

Disconnecting the Mill for Faults in the Power System

The frequency of major voltage disturbances originating in a mill's principal circuits is usually low as compared with the utility.[10] One specific case studied gives a ratio of 27 to 1 over a 23-year period. This is understandable considering the simplicity and very limited extent of the mill's protected distribution system as compared with the extensive network of exposed long lines and diversity of loads of the utility. Heavy faults at considerable distance may result in violent voltage and power fluctuations at the mill. If the effects are not isolated or minimized, mill processes will be interfered with. Such disturbances cause the following equipment and process outages:

1. Loss of the complete powerhouse.
2. Loss of powerhouse auxiliaries.
3. Breakage of the sheet in the paper machines.
4. Loss of the paper machine auxiliaries.
5. Loss of continuous process drives.

There appears to be little reason for a mill to attempt to operate through such voltage and power disturbances if it has in operation sufficient generating capacity to carry the essential load. By means of high-speed, fault-detecting relays and modern circuit breakers it is possible to isolate the mill on its own generators very quickly by opening the tie-line circuit breaker. If the disconnection is effected sufficiently fast, mill processes are not seriously affected. Some voltage and frequency variations will occur while the turbine generators stabilize under the suddenly applied new load conditions. The total fault removal time should be made as short as practicable. Values of 7 to 10 cycles can be realized with present-day standard switchgear, at 15 kv and below. The fault-detecting relays will be connected to instrument transformers of the tie circuit. If the power delivery voltage is above 15 kv, faster speeds of tie interruption can be realized by opening the transformer low-voltage seconday breaker instead of the more heavily constructed primary breaker. See Figs. 1 and 2.

If the generating units in operation are not capable of assuming the increased load, nonessential load feeders can be automatically dropped, simultaneously with detection of the fault condition. This requires that the mill load be departmentalized with a sufficient number of feeders. The required results are obtained most easily and simply with the simple radial type of power distribution system. The circuits should provide for maintenance of lighting in all departments during the emergency.

High-speed (calibrated in cycles), nondirectional, induction-type overcurrent relays are adequate for the usual application. These relays are easily adjusted to start operating on any desired value of overcurrent. By means of a lever, the operating time may be adjusted to any required value. For most applications lever settings of one-half division or less should be used to obtain maximum speeds of tie severance. The relays can be substituted in place of standard induction overcurrent types. Initial settings are established by calculation, followed by experimentation as dictated by the results obtained.

Low-energy directional overcurrent and power relays may be justified in specific applications, with suitable provisions made for the mill to deliver real and/or reactive power to the utility. Since major faults within the mill occur so infrequently the gains possible by directional relaying may be limited to special cases.

For best results the turbine governors and extraction mechanisms must operate fast and with minimum tendency to hunt when load is suddenly applied. Generator voltage regulators should also be fast-acting with their cross-current compensating circuits adjusted so that all machines properly divide the reactive load.

After best methods and adjustments have been established for severing the tie automatically, a few circuit-controlling devices in the mill may still malfunction under the worst fault conditions. These should be handled individually using applicable correctives as shunt trip attachments, lower voltage operating coils, time-delay undervoltage attachments, latched-in devices, etc.

If the speed-regulating turbine generator (for isolated mill operation) is equipped with an isochronous speed governor, isochronous speed regulation of mill frequency may be automatically re-established when the tie is interrupted to maintain electric clocks essentially on time. Isochronous operation must be automatically discontinued at the instant the tie is re-established to prevent violent load changes on the turbine for slight changes in utility frequency.

Stabilizing Mill Generators by Carrying Reactive Load

Usual practice is to allow the utility to carry some portion of the mill's reactive load, as determined by the contract power factor. This may allow the mill to carry more of its own kw load more efficiently with fewer generators on the bus.

However, if power system fault conditions require rather frequent opening of the tie, improved mill voltage stability can be obtained by increasing the normal lagging reactive load carried by its generators, even to the extent of increasing the number in operation. Such operation may be extended advantageously to where the mill delivers reactive power to the utility while simultaneously drawing real power from it.

The increased excitation required establishes higher orders of flux in the generators whose voltage is consumed in generating the lagging current. At the time of fault removal the excess excitation exists to assist in picking up the mill's total kilovolt-ampere load with minimum disturbance and drop in bus voltage. The rate at which the voltage recovers is also improved. The amount of overexcitation required for most satisfactory results depends on the average kw load the generators are carrying prior to the fault, and on the power factor of the mill load to be assumed. An approximate maximum excitation value would be that corresponding to the total kw and kilovar load to be assumed.

Fig. 11 partially illustrates the increased stability obtained for assumed initial and final generator kilovolt-ampere loadings of 60 and 96 per cent. In the first case the utility supplies all the mill reactive with the generators operating at unity power factor. In the second case the generators are carrying the mill reactive plus a portion of the system reactive at 62.5-per-cent power factor. The improved voltage stability is indicated by the relative lengths of phasors E_1 and E_2. In both cases the mill load is 100-per-cent kw, 125-per-cent kilovolt-amperes at 80-per-cent power factor in per cent of generator rating. The usual mill will have better than 80-per-cent power factor.

When new turbine generators are to be purchased, consideration should be given to ordering an oversized (next higher rating) generator. Standard industrial turbine-generator units (10,000 kw and smaller) are rated 100-per-cent kw at 80-per-cent power factor, of which the turbine is capable of driving the generator at its full kilovolt-ampere rating at 100-per-cent power factor or 125-per-cent kw. The turbine is therefore essentially capable of supplying the full power needs of the next larger generator for operation at 100-per-cent kw load and 80-per-cent power factor.

Oversized generators powered by normal capacity turbines still further improve mill voltage stability when the power tie is broken. The oversized generator may make it possible to operate with fewer generating units in operation and at considerable saving in fuel. Under certain operating conditions extraction and back-pressure turbines in pulp mill process may be capable of exceeding the normal guaranteed kw output capability, which may still further justify the increased cost of an oversized generator.

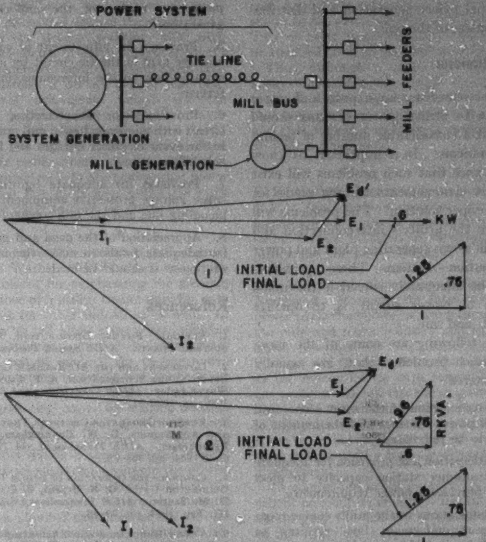


Fig. 11. Improved mill voltage stability obtained at the instant of power-tie interruption when mill generators are normally used to carry substantial amounts of reactive load

①—Low degree of stability obtained with no initial reactive load
②—High degree of stability obtained with 75-per-cent initial reactive load

Circuit Breaker Versus Fused Power Tie

When the supply voltage is above 15 kv the cost to provide for a primary circuit breaker introduces the lower cost alternate of high-voltage power fuses. Fuses have the single advantage of lower first cost. However, they possess decided disadvantages:

1. A blown fuse subjects the entire mill to single-phase operation which endangers operation of the mass of electric equipment.
2. A fused tie requires that means be provided to detect blown-fuse operation and open a principal secondary circuit breaker to relieve the mill from the hazards of single-phase operation.
3. If the mill has generators of limited capacity, detection of single-phase operation may be difficult and unreliable.
4. Continued unbalanced load current operation resulting from a blown fuse may dangerously overheat the mill's turbine generators.[11]
5. The detection and isolation of a retained single-phase short circuit resulting from improper fuse operation can in a matter of a few seconds seriously damage the rotors of any of the mill's generators subjected to the condition.[12]
6. The blowing of prime circuit fuses constitutes a nuisance to mill operation which is usually solved by installing oversize fuse elements without concern for the loss in equipment protection that may result. The condition is more serious for mills having operating generators.
7. The restoration of service will require longer time and be more hazardous, particularly during storms or at night.
8. The engine room operator is deprived of direct control of the high-tension circuit in emergencies.

In general, the use of fuses in a major power tie to a mill with generators increases the general level of protection required for all mill equipment. The cost and complication of the additional protective apparatus and the increased maintenance required to insure its operation will seldom balance the anticipated savings.

Many mills have sufficient qualified personnel to maintain the more elaborate systems of equipment protection. However, the system which requires the simplest forms of relay protection with the minimum number of protective devices will contribute more to continuity of service than the system which requires more elaborate forms of protection which are not properly understood or maintained.

Fuses may properly be used in power ties to small mills that do not operate generators in parallel with the system, but their use will always be attended by some of the disadvantages listed.

Lightning and Surge Voltage Protection

The self-contained mill power system without exposed lines may not require lightning protection, particularly if located in a low isokeraunic level region.

A utility power tie imposes the necessity for lightning protection. Lightning and other power system surges will be transmitted to the mill in proportion to the system exposure. Such surges are transmitted even through transformers and are a hazard to apparatus of low voltage-impulse level such as generators, motors, cables, and switchgear. In general, station-type lightning arresters are justified to provide the required degree of protection for the large investment involved.

In addition to lightning arresters for the incoming line, surge protective apparatus[13,14] should be provided on the principal power generation and distribution busses of the mill.

Conclusions

Interconnecting large pulp and paper mills with utility power systems should present a considerable number of technical problems. It is important to know in advance that such problems will exist so that arrangements can be made for their proper solution. The problems will involve the utility's interconnection and the mill's own generating plant and power distribution system. These problems must be adequately solved if the mill's source of power supply is to operate reliably and safe.

The following are some of the more important problems that are usually encountered:

1. Determination of the most economical form of power contract and the amount of power to be purchased.
2. Anticipation and provision for adequate tie transformer station capacity to meet future purchased power requirements.
3. Modifications in the mill's engine-room switchgear and main bus circuitry to provide for its safe operation.
4. Provision for control of purchased power by means of the mill's turbine generators.
5. Provision for adjustment or control of reactive load in the tie circuit by means of suitable taps on the interconnecting transformer.
6. Provisions for disconnecting the tie circuit with minimum loss in mill production in the event of power outages and faults in the utility system.
7. Provision for adequate lightning and surge voltage protective equipment for the incoming line and the mill's generator bus.
8. Appreciation of the need and provision for adequate funds to make the interconnection as it should be made.

References

1. Electric Power Distribution for Industrial Plants. *AIEE Special Publication S-4.*

2. Governors and the AIEE-ASME Governor Performance Specification, **A F. Schwendner, Wayne Astley.** *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 1330–32.

3. Turbine Generator Controls, Protections, and Accessories, **G. W. Cunningham, M. A. Eggenberger.** *AIEE Transactions*, vol. 73, pt. II, June 1954, pp. 455–65.

4. Controls for Operation of Steam Turbine-Generator Units, **O. N. Bryant, C. C. Sterrett, D. M. Sauter.** *AIEE Transactions*, vol. 73, pt. III, Feb. 1954, pp. 79–88.

5. A New Regulator and Excitation System, **J. T. Carleton, P. O. Bobo, W. F. Horton.** *AIEE Transactions*, vol. 72, pt. III, Apr. 1953, pp. 175–83.

6. Bibliography of Magnetic Amplifier Devices and the Saturable Reactor Art, **James G. Miles.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2104–23.

7. Magnetic Amplifiers in Industry, **F. N. McClure.** *Westinghouse Engineer*, East Pittsburgh, Pa., vol. 10, Sept. 1950, pp. 201–05.

8. Industrial Application of Rototrol Regulators, **W. R. Harris.** *AIEE Transactions (Electrical Engineering)*, vol. 65, Mar. 1946, pp. 118–23.

9. Rototrol Excitation Systems, **J. E. Barkle, C. E. Valentine.** *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 529–34.

10. Electric Power in the Pulp and Paper Industry. *AIEE Special Publication S-19*, January 1948.

11. Effects of Negative-Sequence Currents on Turbine-Generator Rotors, **E. I. Pollard.** *AIEE Transactions*, vol. 72, pt. III, June 1953, pp. 404–06.

12. Generator Negative-Sequence Currents for Line-to-Line Faults, **R. F. Lawrence, R. W. Ferguson.** *Ibid.*, pp. 9–16.

13. Survey of Lightning Protective Equipment for Rotating A-C Machines, **AIEE Committee Report.** *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 516–19.

14. Lightning Protection for Rotating Machines, **G. D. McCann, E. Beck, L. A. Finzi.** *AIEE Transactions*, vol. 63, 1944, pp. 319–33.

15. Protection of Generators Against Unbalanced Currents, **J. E. Barkle, W. E. Glassburn.** *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 282–86.

16. Protection of an Industrial Plant Connected to a Utility System with Reclosing Breakers, **D. F. Langenwalter.** *Proceedings*, Technical Association of Paper and Pulp Industry, New York, N. Y., vol. 36, no. 5, May 1953, pp. 199–203.

Discussion

F. Alexander (H. A. Simons Ltd., Vancouver, B. C., Canada): This excellent paper is a clear and concise presentation of a problem which is of growing importance to the pulp and paper industry. The authors deserve the thanks of all busy engineers, who all too often lack time to search the scattered literature and assemble a comprehensive picture of the problem.

The first section of the paper ably covers all the salient points of the general problem and leaves little area for comment except for minor details which may be put forward as further amplification rather than modification. Several such items occur to me and are advanced to point up certain features which may require special attention in certain cases.

Opening the Utility Tie Under Both Normal and Abnormal Conditions

Assuming that the tie breaker used for synchronizing is located on the load side of the tie transformer, means must be provided to positively open this breaker whenever the utility supply is interrupted by the operation of protective equipment on the utility system other than the tie breaker. For example, should a utility breaker remote from the mill open, it should be prevented from reclosing until the synchronizing tie breaker is opened. One way in which this can be accomplished is to supervise the utility line on the load side of such a breaker using a voltage sensitive relay arranged to prevent either manual or time delay automatic reclosing until the tie breaker is opened and the intervening line is at zero voltage.

Reverse Current Protection

In some cases it may be desirable to prevent any reverse power flow, particularly if the mill generating capacity is substantially less than the normal load. As pointed out by the authors, sensitive directional relays can be used and where these can be made to operate on the magnetizing current of the tie transformer it will assist in opening the tie breaker if the utility supply is interrupted. Relays of this type would be too sensitive for normal switching surges due to synchronizing but the tie breaker control switch can be arranged with contacts to bypass these relays during switching operations.

Dropping Out Nonessential Load

It is not always desirable to automatically drop out a fixed number of feeder breakers when the utility supply fails. Some flexibility is desirable so as to drop out the minimum number of feeders required to reduce the load to the capacity of the turbines. The number to be dropped would depend on the mill load. One method of accomplishing this is to back up the sensitive reverse power relays with a frequency sensitive relay arranged in conjunction with time delay relays so as to progressively drop out feeders until a reasonable system frequency is restored.

Turbine Governing

In the case of a mill power plant consisting of noncondensing turbines only, such units would normally operate under pressure governing based on maintaining relatively constant extraction and back pressures with the electric output subordinated to process steam requirements. System frequency would be maintained by the utility. When the utility supply is interrupted means should be provided to automatically transfer the mill turbine from pressure governing to speed governing so as to maintain the frequency and maximum electric output. This will generally involve blowing exhaust steam to atmosphere and is not particularly objectionable unless the outages are frequent or of long duration. The turbine would be transferred back to pressure governing when normal system conditions are restored and the mill turbine again synchronized with the utility system.

E. E. Dale (Crown Zellerbach Corporation, Seattle, Wash.): The authors have done an excellent job in putting down in comprehensive form the problems encountered in interconnecting the electric power system of a large paper mill with that of a large utility. This paper could very well serve as a guide to an engineering staff in setting up a study prior to making such a tie.

To add emphasis to one of the points outlined by the authors under "Power-System Layout and Switchgear Requirements," I should like to add this information gained through experience. In 1948 the Bonneville Power Administration ran a single-circuit 115-kv line into the Port Angeles area to serve the existing power system. This system consisted mainly of two hydrogenerating stations, with a total of 26,000 kva of installed capacity, two large paper mills, each with steam generation, and a tie to an existing utility. The additional fault contribution from the Bonneville Power Administration system, and the low rate of fault decay, made it necessary for the company to replace 28 large circuit breakers both in 66- and 2.4-kv circuits due to low interrupting capacity.

This experience has made us somewhat cautious. In enlarging the tie to a utility at another of our mills undergoing an expansion program, we investigated the premium one would have to pay for a 500-megavolt-ampere interrupting-capacity 1,200-ampere air circuit breaker over a 250-megavolt-ampere interrupting-capacity breaker at 15 kv. The premium was $200.00 per circuit. We considered this additional cost a justifiable expenditure, not knowing positively just what a fast-growing mill would require in power development.

A. M. Buck (Weyerhaeuser Timber Company, Everett, Wash.): This is a timely paper and the authors are to be congratulated for a thorough coverage of the subject. While it deals with the pulp and paper industry, it could also point the way to some clearer thinking in other industries.

The paper covers the basic facts and requirements of an industry that is fast loosing its low-cost fuel supply. It emphasizes the point that, when power ties are made with utilities, adequate capacity should be installed originally, at least provision should be made for easy expansion, otherwise production may be lost at a critical time later on. As intimated, process steam requirements are large, and heat balance is critical, requiring a stable utility tie and smooth control of utility and generated power. Hopes of an ideal utility voltage are not always realized, owing to the utilities' load factor, so perhaps the paragraph dealing with automatic or at least manual underload tap changing should receive some thought as the exchange of reactive power may be a problem at times. Usually it is economical to operate the back pressure units at full load using the controlled reducing valve to take care of fluctuations.

I believe the simplest way is to allow the utility to take care of all load swings, providing it is possible to negotiate a fair contract and providing the peak penalty is not too high and the load is such that a reasonable amount of that block of power can be used. As is suggested, a reliable type of load control may be indicated. With the large present-day chippers, peaks have become more critical, especially in mills that have no grinders.

As is stated, electric energy converted to thermal energy is definitely not economical, although some of the mills have experimented with its use as an auxiliary aid and have found that the results could not justify the cost.

If it is feasible as suggested, it is advisable to cut loose from the utility, should it get into trouble, as quickly as possible. If the control on the condensing turbine is fast enough it will save a lot of stock from going to the sewer, and if the mill lends itself to some load control the sheet can be kept on the drying machine, this with the judicious use of delay voltage trips and motor-generator sets on maintained contact controls. I do not believe that nondirectional overcurrent relays are the answer in most cases. I would favor frequency or directional power relays. I think the authors should have stressed adequate power circuit breaker application a little more, as I do not believe that power fuses have any place in a utility tie when they are likely to blow on power exchange.

Fred Thompson (General Electric Company, Seattle, Wash.): The authors are to be congratulated for emphasizing the growing importance of a utility tie not only for pulp mills but for other industries as well. In addition to the lower power costs possible with a properly controlled tie, even greater savings can be made if the tie-line power rates are set up to reflect the benefits that accrue to the utility. Where a utility is also purchasing power on a block demand basis from a power pool, a relatively small amount of power supplied by the mill generators during a short interval of peak demand can mean real savings to the utility. This saving can be balanced against the increased kw demand required by the mill during a planned shutdown of one turbine unit for maintenance. A mutually advantageous power contract can then be agreed upon.

Automatic control of both the tie-line kw and kilovars is almost always required to take advantage of the usual rate schedule. When large fluctuating loads, such as large pulp mill chipper drives, are encountered, the high response of the amplidyne load regulator shown in Fig. 5 will be needed. For the usual installation of slowly changing loads an impulse-type controller is satisfactory.

There seems to be a mistake in the section entitled "Disconnecting the Mill for Faults in the Power System," since the mill should definitely try to ride through the load disturbances listed. In many installations additional protection other than high-speed overcurrent relays will be required. Nondirectional overcurrent relays must be set fairly high in order to ride over switching and synchronizing surges whereas directional power relays can be set low since power flow will normally be from the utility to the mill.

If the utility breaker is of the reclosing type, some method must be provided for holding off the reclosing when the outage time is long enough for the mill synchronous machines to get out of step with the utility system. As shown in reference 16 of the paper, slow reclosers are of little value since they must be prevented from reclosing. Fast reclosers can be used but the plant loads must be examined carefully to be sure satisfactory reclosing can be obtained.

One method of relieving the excess load on the mill generators resulting from utility outages is to drop nonessential load when the tie-line breaker opens. Under certain load conditions, this may not be necessary and under other conditions additional load may need to be dropped. Underfrequency relays connected to the mill generator bus will provide for these various conditions and drop load only when necessary, as shown by the slowdown of the mill generator. Two or more underfrequency relays may be required to drop the load in steps.

Thus, although a utility tie is almost always advantageous, each installation must be thoroughly examined to assure that proper kw and kilovar control is applied and that sufficient relaying is provided for all system conditions.

E. K. Murphy (Rayonier, Inc., New York, N.Y.): The authors deserve great commendation for a precise presentation of the many factors involved in interconnecting utility and mill power systems.

The economic considerations outlined have been based on conditions in the Pacific Northwest. However, the ratio of electric power to thermal power is rising throughout the pulp and paper industry and the principles described by Mr. Rose and Mr. Springer will therefore be applied elsewhere as the economics for each respective area are evolved. Given the simultaneous conditions of a mill on condensing operation and a utility with available off-peak power, it is fast becoming inevitable that the organizations will negotiate a mutually advantageous arrangement.

As pointed out in the paper, there are excellent possibilities of improving mill system stability by the delivery of reactive power to the utility while simultaneously drawing real power from the utility. This is a somewhat common practice in some utility interconnections but it has not been common in the type of tie under discussion. It is believed that more of this type of operation will be found in the future as the mutual advantages become more generally realized. Greater impetus will be given to moves in this direction as industry recognizes that the investment for this feature will pay off as the utilities recognize that a rate credit may be in order.

The paper is quite clear on the point that there is little reason for a mill to attempt to operate through utility system disturbances; further, that speedy isolation, coupled with ability to pick up mill load, is quite necessary. There is now available an inertia relay which lends itself very well as one of the devices to maintain a mill in operation in the event that such disturbances occur. This relay recognizes the rate of change of frequency and operates whenever frequency changes at a rate in excess of a preselected value. Such a relay can differentiate between the type of frequency change which spells shutdown if allowed to continue and the type of change that can be tolerated or handled in the normal manner. The relay may be applied to initiate the tripping of nonessential circuits in sequence, the sequence being followed only as long as the rate of change of frequency is too high. It may also be desirable to make the utility tie breaker the first element in such a sequence. This approach drops the minimum number of circuits.

H. D. Hunkins (U. S. Bureau of Reclamation, Denver, Colo): The authors are to be congratulated for their thorough discussion of the subject. It is interesting to note how closely the problems and solutions compare with those involving the interconnection of a power system.

If the pulp mill were located in an area remote from power company generators, could energy costs be reduced on the basis of providing stand-by capacity for a small portion of the companies' local loads or by absorbing line charging vars from the company system during light load periods?

Under the section "control of Reactive Power in the Tie," four methods are given for increasing var supply at the mill. It would appear that static capacitors might be added to the list of methods for supplying reactive power during heavy load periods. It might be best to place a part of these at the ends of plant feeders to reduce currents.

Under the section "Regulators for Controlling Real and Reactive Load in the Tie Circuit," it is suggested that a kw-hour integrator be superimposed on the constant kw regulator to hold the net input energy to a fixed quantity over a given time. If the power inputs which exceed the contract demand last long enough to require this energy integrator, is it not likely that an integrating demand meter usually used on industrial customer connections would show the increased demand resulting in an increase in the billing?

W. Kullrich (Rayonier, Inc., Shelton, Wash.): In this paper the authors have stressed the use of full-load tap changers on utility tie transformers, particularly where the control of reactive power is necessary and where this control is difficult to handle without tap changing. I would like to cite a specific example of a case in my experience where abnormal ability voltage variations over a period of months created a serious and costly condition in one mill.

In this case a system was planned for a nominal voltage of 12,470 volts and no-load taps were provided for normal variation from this value. During the first few months of operation, the voltage varied from slightly under 12,470 to 13,600 which resulted in over 500 volts on all 480-volt services on the mill system at times. No means except capacitors were available for control of reactive flow to or from the utility and with the mill on 24-hour operation continuously no taps could be changed without a plant shutdown. The power factor penalty paid by the mill was from $600 to $800 per month for several months. This condition was the result of the utility attempting to meet a low-voltage condition in one plant and creating an undesirable high-voltage condition in others and through no fault of their own.

Naturally, the proper solution would be for all plants operating on the utility system to provide the proper taps for operation at the nominal voltage established by the utility for the area, but this instance illustrates conditions which could be encountered in a mill when interconnecting with a utility system. During this high-voltage condition the control of the reactive flow could not have been controlled by mill generators or capacitors since the resultant mill voltage was already too high and reducing the reactive flow to the mill through such means would aggravate the voltage condition.

The condition was later corrected and stable conditions established, but before the mill could be shut down and the no-load tap changers on all mill transformers reset a period of abnormally low voltage was experienced.

H. A. Rose and **H. E. Springer:** The authors appreciate the many constructive comments of the discussers. These constitute valuable supplemental information to the principal subjects of the paper.

The four important points concerning tie-line power control discussed by Mr. Alexander should be given consideration for all mill and utility power system interconnections. Such considerations if made prior to purchase of the tie equipment will insure the most practical and economical solutions to the problems and will avoid the expense of making modifications later.

Mr. Alexander points out the practicality and limitations of blowing exhaust steam to atmosphere with noncondensing turbines when the power tie is interrupted. If many interruptions are anticipated, the additional cost for a condensing unit may be justified to avoid the noise and nuisance factors usually associated with atmospheric exhaust operation. A detailed study can be made to determine the probable number of tie-line interruptions per year that the mill will have to make to avoid process interruptions. The paper indicates that this figure can be considerably larger than a preliminary investigation might disclose.

Mr. Dale's discussion also emphasizes the importance of an adequate engineering analysis before purchasing equipment for a power tie. A few dollars wisely spent initially can save many later on through adequate provisions to meet the mill's future growth requirements. The plans for a major utility interconnection should usually anticipate the future addition of at least one turbine-generator unit.

Mr. Buck mentions the desirability of the utility taking the load fluctuations of the mill. Assume that power is purchased on a 15-minute or 30-minute demand basis. The swinging loads of wood mill operations will average a median value for the demand period. The load controlling turbine generator can therefore be arranged to regulate for essentially constant purchased power as measured at the end of each demand interval, but with the load swings being carried by the utility. As mentioned in the paper, a load integrator would be desirable to obtain minimum cost if power is purchased on a block rate demand schedule.

Most mills actually use nondirectional overcurrent relays. This is because their normal operating load is reasonably constant and not subject to heavy load swings such as result from starting and operation of large motors in the order of 1,500 horsepower. It is therefore usually possible to set nondirectional overcurrent relays to trip high-speed on utility faults approximating 150 per cent of mill load. When heavy load swings are involved such as for starting and operation of large chippers, grinders, and hydraulic barking pump motors, this may no longer be possible. In these cases, directional-type relays may be required. The application of directional power relays should always give consideration to the possible need and provisions for supplying real power to the utility.

Mr. Thompson expresses the opinion that mills should attempt to ride through power system disturbances. We agree with this concept as a general principle. However, our experience indicates that for best results it is necessary that the protective relay system be selectively adjusted to trip on a considerable number of system disturbances that would not have caused equipment outages in order to insure a satisfactory actual minimum of such outages. This is the intended philosophy on which the paper discusses automatic tie-line disconnecting.

In general, the reversible d-c governor-head motor of Fig. 5 may be equally well controlled by any of the five types of load regulators enumerated in the paper, the type required being dependent on the degree of accuracy and speed of response required. Until recently, rotating regulators such as Rototrols were probably the most practical and reliable types of high-speed regulators for mill services. The new magnetic-amplifier types of static regulators now show promise of superseding rotating regulators for most regulating applications.

Mr. Murphy emphasizes a mill's kw and var load conditions that can benefit a utility. In this connection, we again emphasize the importance of considering the application of oversize rated generators to be driven by normal rated extraction and condensing turbines for mill service.

Mr. Hunkins suggests the use of static capacitors for control of mill reactive. Such arrangements may be of particular importance to small mills and to large ones that do not operate electric generating plants. Large-size bus capacitor banks should in general be sectionalized with appropriate circuit breakers and control relays to sequentially reduced connected var capacity to avoid overvoltages on mill equipment during periods of light loading of the mill and/or the utility system.

We believe that a study would be required to determine if the utility could better afford to operate a remote located power plant to supply the unbalanced portion of the mill's power needs. If the local utility loads were large in proportion to the mill's needs, it would probably be more economical for the mill to provide for all of its power requirements even to the extent of operating isolated.

If joint operation was mutually agreed to, the mill would most likely be an electrical stabilizing asset to the isolated community. A synchronous condenser or sectionalized capacitor bank might be needed to avoid excess var loading of the mill's generators.

The kw-hour integrator referred to in the paper recalibrates the kw tie-line regulator according to the actual rate of energy flow so that at the end of each demand period essentially the allowable number of kw-hours are taken. The tendency for the integrator to accumulate a negative deficiency in consumed kw-hours is corrected within each demand period to the sensitivity adjustment of the integrator.

Mr. Kullrich's experiences emphasize the importance of providing the tie transformer with an adequate range of voltage equalizing primary winding taps. If the tap range is large and requires equalizing operation at low delivery voltage, full kva capacity taps will be required. The tie transformer should be equipped with a no-load type of tap changer for external manual operation.

A Generalized Method for Determining the Closed-Loop Frequency Response of Nonlinear Systems

LUTHER T. PRINCE, JR.
ASSOCIATE MEMBER AIEE

Synopsis: One of the formidable problems in the area of feedback control concerns the analysis and synthesis of nonlinear systems. Of the several methods available for treating nonlinear systems, one of the more widely used is the sinusoidal response method. Valuable information about the characteristics of nonlinear systems can be obtained from the open-loop frequency response, but the full potential of this method has not been exploited because apparently there is no widely accepted procedure for determining the closed-loop frequency response. This paper describes a method whereby the closed-loop frequency response of an important class of nonlinear systems can be determined easily. Development of this method is dependent upon representing the nonlinear element as a so-called equivalent gain. The method is described and illustrated by analytically determining the closed-loop frequency response of an acceleration-limited system and an error-limited system. The results of the use of this method compare favorably with solutions of the two problems as obtained from an analogue computer. A few of the nonlinear phenomena which can be reasonably approximated by this method are saturation, limiting, backlash, and dead space.

SINUSOIDAL analysis of nonlinear feedback control systems is an outgrowth of valuable techniques already developed in studies of linear feedback control systems. Since nonlinear behavior cannot be described by ordinary linear differential equations with real, constant coefficients, the powerful tool of superposition is not applicable; consequently, a frequency-response analysis is not as meaningful. Nevertheless, considerable information about nonlinear systems can be obtained through a sinusoidal analysis, as exemplified by the work of Kochenburger,[1,2] Johnson[3] and others.

The most widely used method is based on a representation of nonlinear elements by describing functions which facilitate analysis and synthesis using the open-loop frequency response. By using describing functions, it is possible to determine whether or not there is danger of sustained oscillations or destructive instability and how such behavior can be avoided. Unfortunately, there seems to have been no general procedure developed for determining the closed-loop frequency response: hence, quantitative information other than stability considerations cannot be generally obtained for closed-loop operation. By approximating nonlinear elements by equivalent gains, it is possible to simplify the mathematical equations so that the closed-loop frequency response of many nonlinear systems can be determined easily.

The sinusoidal response method is a special technique which is useful in studying the behavior of a certain class of nonlinear systems, and determination of the closed-loop frequency response must be considered as a supplement to existing sinusoidal techniques. The value of the method can be determined only by its applicability to specific problems.

The Equivalent Gain Approximation

The sinusoidal response method of analyzing nonlinear systems is dependent generally on the validity of two assumptions: 1. The harmonics at the output of the nonlinear element can be neglected because of the low-pass filter characteristics of servo systems. 2. The input to the nonlinear element can be approximated by a sinusoid. On the basis of these assumptions, it is customary to represent the nonlinear element by an approximate transfer function in order to use operational notation throughout the system. Many of the techniques developed in linear analysis can then be extended to nonlinear systems.

The equivalent-gain representation is similar to the describing function, but possesses certain advantages for closed-loop analysis. The describing function is defined as the complex ratio of the fundamental harmonic component of the output to the sinusoidal input of the nonlinear function. The complex ratio of the maximum value of the output to the maximum value of the input to the nonlinear element characterizes the equivalent-gain approximation. The distinction between the terms "nonlinear function" and "nonlinear element" will be that the former may include the latter plus linear elements. To simplify the nonlinear equations, it is customary to replace the true characteristics of the nonlinear element by a piecewise linear approximation. This simplification is usually made in order to determine a simple analytical expression relating the variables of the nonlinear device. When used in a sinusoidal analysis, the equivalent-gain definition implies the existence of a fictitious sinusoidal variation of the output of the nonlinear element which has the same period as the input and the same peak amplitude as the true output variation of the nonlinear element, as shown in Fig. 1. A rigorous analytical justification of this definition is beyond the scope of this paper; however, analogue-computer solutions have been obtained to serve as a check on the error introduced by this approximation. It will now be shown how the closed-loop frequency response of a nonlinear system can be determined by representing the nonlinear element with an equivalent-gain function.

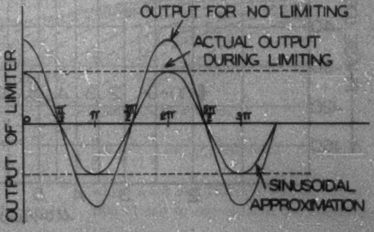


Fig. 1. Wave forms at the output of a limiter

Paper 54-280, recommended by the AIEE Feedback Control Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21-25, 1954. Manuscript submitted March 22, 1954; made available for printing April 23, 1954.

Luther T. Prince, Jr., is with the Minneapolis-Honeywell Regulator Company, Minneapolis, Minn.

This work was sponsored in part by the United States Air Force (WADC) under contract *AF33-(038)-22893*, with Dr. C. L. Morrison of the Aeronautical Research Laboratory as project engineer.

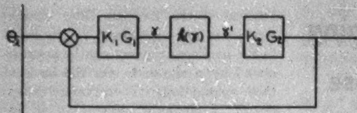


Fig. 2. Block diagram of a nonlinear feedback control system

Nomenclature

K = gain constants
$G(s)$ = frequency variant part of linear transfer functions
$k(\gamma)$ = equivalent gain
γ = input to nonlinear element
γ' = output of nonlinear element
θ_i = input forcing function
θ_0 = output of feedback system
$\dot{\theta}_0$ = first derivative of the system output
$\ddot{\theta}_0$ = second derivative of the system output
S = Laplace operator
$f[k(\gamma)]$ = equivalent-gain function
$\ddot{\Theta}_0$ = maximum possible value of $\ddot{\theta}_0$
$j=\sqrt{-1}$
ω = angular frequency in radians per second
ω_{n0} = system natural frequency during linear operation
ζ_0 = system damping ratio during linear operation
u = nondimensional angular frequency (ω/ω_{n_0})
T_m = motor time constant

The Closed-Loop Frequency Response

Fig. 2 is a block diagram of a feedback control system containing a nonlinear element. The equivalent gain of the nonlinear element is represented by $k(\gamma)$. As contrasted to a describing function, which can be both frequency dependent and amplitude dependent, an equivalent gain, is dependent only on the amplitude of the input signal if the static characteristics of the nonlinearity equally describe the dynamic operation, i.e., no energy storage in the nonlinear element. More will be said about this later. If a fixed sinusoidal signal, θ_i, is introduced at the input of the system shown in Fig. 2, γ will assume a sinusoidal steady-state value after all transients have died out; consequently, $k(\gamma)$ will be constant. The equivalent gain can be determined from the following relationship

$$k(\gamma)=\gamma'/\gamma \tag{1}$$

since γ' can be determined from the static characteristics of the nonlinear element for all values of γ. If, however, the magnitude or frequency of θ_i is changed, $k(\gamma)$ will assume a new value since the magnitude of γ reflects any change in θ_i. Hence, $k(\gamma)$ can be considered a quasiconstant gain, and the nonlinear element can be represented by an amplitude-dependent transfer function. The closed-loop frequency response of the configuration shown in Fig. 2 is determined by the following equation

$$\frac{\theta_0}{\theta_i}(s)=\frac{k(\gamma)K_1K_2G_1(s)G_2(s)}{1+k(\gamma)K_1K_2G_1(s)G_2(s)} \tag{2}$$

where $S=j\omega$, but the solution of this equation is indeterminable because θ_0 and $k(\gamma)$ are both unknown. Since $k(\gamma)$ can be represented as a quasiconstant gain under steady-state conditions, a family of closed-loop frequency-response curves can be computed with $k(\gamma)$ as a parameter. These loci can be determined from the following equation

$$\frac{\theta_0}{\theta_i}[S,k_n(\gamma)]=\frac{k_n(\gamma)K_1K_2G_1(s)G_2(s)}{1+k_n(\gamma)K_1K_2G_1(s)G_2(s)} \tag{3}$$

where the subscript n denotes the value of $k(\gamma)$ for each curve, and $S=j\omega$. Such a family of curves is shown in Fig. 3 with both magnitude and phase plotted. If a

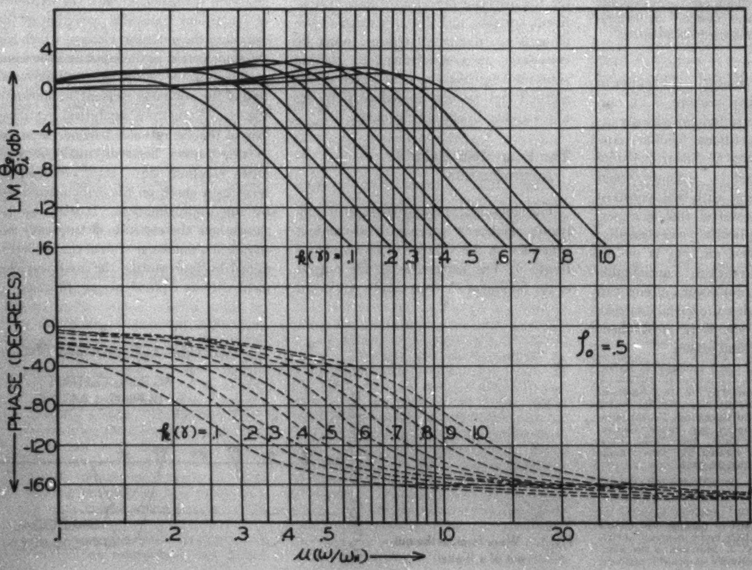


Fig. 3. Closed-loop frequency response curves with k(γ) as a parameter

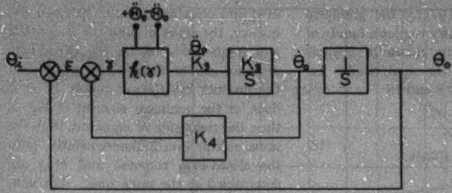


Fig. 4 (left). An acceleration-limited system

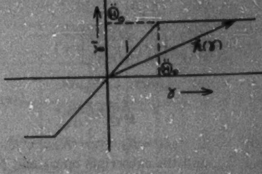


Fig. 5 (right). Static characteristics of a limiter

sufficient number of these curves is plotted in the range of possible values of $k(\gamma)$, then points on the closed-loop frequency response of the nonlinear system will occur somewhere along the constant $k(\gamma)$ curves for different values of the magnitude and frequency of θ_i. These points can be located exactly if $k(\gamma)$ can be determined as a function of the magnitude and frequency of the forcing function θ_i during closed-loop operation. The method presented in this paper makes it possible to determine $k(\gamma)$ from equations 1 and 3 with the use of the block diagram of Fig. 2.

Analytical Determination of $k(\gamma)$ in a Closed-Loop System

By using the approximate static characteristics of the nonlinear element, it is possible to derive an analytical expression for $k(\gamma)$ as some function γ multiplied by a constant. Conversely, γ can be determined as the product of a constant times some function of $k(\gamma)$; consequently, either $k(\gamma)$ or γ can be considered as the independent variable. This is the key to the solution of the nonlinear equation. It is possible to obtain two equations in which $k(\gamma)$ and γ are the only unknowns. Since $k(\gamma)$ can be considered as an independent variable, it is possible to solve these equations simultaneously, and to eliminate γ, and thus determine $k(\gamma)$ as a function of the magnitude and frequency of the forcing function. Having determined $k(\gamma)$, the closed-loop frequency response can be obtained from the families of curves computed from equation 3.

With the block diagram of Fig. 2, it is possible to determine the transfer function which relates γ to the forcing function θ_i. This equation is obtained easily from the transfer function, $(\theta_0/\theta_i)(s)$. The relationship is

$$\frac{\gamma}{\theta_i}(s)=\left[\frac{\theta_0}{\theta_i}(s)\right]\frac{1}{k(\gamma)K_2G_2(s)} \tag{4}$$

where $(\theta_0/\theta_i)(s)$ is defined in equation 2.

By rearranging equation 1 and dividing both sides by θ_i, another equation relating γ to θ_i results

$$\frac{\gamma}{\theta_i}=\left[\frac{\gamma'}{k(\gamma)}\right]\frac{1}{\theta_i} \tag{5}$$

This equation, obtained from the static characteristics of the nonlinear element, relates only the magnitudes of the variables. Two equations have now been obtained relating the two unknowns, $k(\gamma)$ and γ, to θ_i. By equating the magnitudes of these equations γ can be eliminated, leaving one equation with $k(\gamma)$ as the only unknown.

$$\frac{1}{\theta_i}\left[\frac{\gamma'}{k(\gamma)}\right]=\left|\frac{1}{k(\gamma)K_2G_2(s)}\left[\frac{\theta_0}{\theta_i}(s)\right]\right| \tag{6}$$

At this point it is convenient to make the following substitutions

$$\frac{\gamma'}{k(\gamma)}=\alpha f[k(\gamma)] \tag{7}$$

where α is a constant and $f[k(\gamma)]$, some function of $k(\gamma)$, will be referred to as the equivalent-gain function. Values of $k(\gamma)$ will now be determined which satisfy equation 6, subject to changes in the magnitude and frequency of θ_i. Since equation 6 contains a variable coefficient, $k(\gamma)$, it is expedient to separate the functions containing $k(\gamma)$ from the conventional linear transfer functions as in equation 8.

$$\left|\left[\frac{\alpha}{\theta_i}\right]K_2G_2(s)\right|=\left|\left[\frac{\theta_0}{\theta_i}(s)\right]\frac{1}{k(\gamma)f[k(\gamma)]}\right| \tag{8}$$

Two methods suggest a means of solving this nonlinear equation. If values of $k(\gamma)$ are assumed and the trial value is corrected by noting the magnitude and sign of the resulting error, values of $k(\gamma)$ can be determined which satisfy equation 8. This method usually converges but must be repeated for each change in θ_i. A more direct solution may be obtained by a graphical method. First, equation 8 is rewritten so that curves for several values of $k(\gamma)$ can be plotted.

$$\left|\left[\frac{\alpha}{\theta_i}\right]K_2G_2(s)\right|=\left|\frac{\theta_0}{\theta_i}[s,k_n(\gamma)]\frac{1}{k_n(\gamma)f[k_n(\gamma)]}\right| \tag{9}$$

Both sides of equation 9 will now be separately plotted on the same graph. The left side will plot as a conventional linear frequency response, and the right side will be a family of curves plotted as a function of frequency with $k(\gamma)$ as a parameter. The product $k(\gamma)f[k(\gamma)]$ can be considered a gain which is constant for a specified value of $k(\gamma)$. The intersections of the linear curve and the family of curves obtained from the right side of equation 9 determine the values of $k(\gamma)$ which satisfy this equation as a function of the magnitude and frequency of θ_i. These values of $k(\gamma)$ can be used to determine the closed-loop response from the families of curves shown in Fig. 3. A specific nonlinear system will now be analyzed by the method outlined.

Acceleration Limiting

Fig. 4 is a block diagram of an acceleration-limited system showing that the nonlinear element is located in a loop which contains an energy storage element. The equivalent-gain representation of the nonlinear element is amplitude dependent only because this function relates only the input and output of the nonlinearity no matter where it occurs in the system. By using equation 4, it is possible to obtain the following function

$$\frac{\gamma}{\theta_i}(s)=\left[\frac{\theta_0}{\theta_i}(s)\right]\frac{s^2}{K_3k(\gamma)} \tag{10}$$

where

$$\frac{\theta_0}{\theta_i}(s)=\frac{1}{\dfrac{s^2}{K_3k(\gamma)}+K_4s+1} \tag{11}$$

The static characteristics of the nonlinear element are shown in Fig. 5, from which analytical expression for $k(\gamma)$ can be obtained

$$k(\gamma)=\begin{cases}1 \text{ for } |\gamma|\leq|\ddot{\theta}_0| \\ \dfrac{\ddot{\theta}_0}{\gamma} \text{ for } |\gamma|\geq|\ddot{\theta}_0|\end{cases} \tag{12}$$

Using equations 5 and 7

$$\frac{\gamma}{\theta_i}=\frac{\ddot{\theta}_0}{k(\gamma)}\frac{1}{\theta_i} \text{ for } |\gamma|\geq|\ddot{\theta}_0| \tag{13}$$

where $\alpha=\ddot{\theta}$ and $f[k(\gamma)]=1/k(\gamma)$. Equation 11 can be expressed in the following nondimensionalized form

$$\frac{\theta_0}{\theta_i}(s)=\frac{1}{\dfrac{s^2}{\omega_{n_0}{}^2k(\gamma)}+\left[\dfrac{2\zeta_0}{\omega_{n_0}}\dfrac{\sqrt{k(\gamma)}}{\sqrt{k(\gamma)}}\right]s+1}$$

$$=\frac{1}{1-\dfrac{u^2}{k(\gamma)}+j[2\zeta_0]u} \qquad (14)$$

with $s=j\omega$. $\omega_{n0}{}^2=K_3$, $\zeta_0=(K_{4/2})\sqrt{K_3}$, $u=\omega/\omega_{n_0}$, and the subscripts on ω_n and ζ identify these two quantities as the natural frequency and damping ratio of the system when $k(\gamma)=1$. By substituting equation 14 into 10, the equivalent of equation 4 is obtained

$$\frac{\gamma}{\theta_i}(s)=-\frac{u^2}{k(\gamma)}\,\frac{1}{1-\dfrac{u^2}{k(\gamma)}+j[2\zeta_0]u} \qquad (15)$$

It is now possible to eliminate γ by equating the magnitudes of equations 13 and 15

$$\left|\frac{\ddot{\theta}_0}{k(\gamma)}\,\frac{1}{\theta_i}\right|=\left|-\frac{u^2}{k(\gamma)}\,\frac{1}{1-\dfrac{u^2}{k(\gamma)}+j[2\zeta_0]u}\right| \qquad (16)$$

which can be rearranged into the same form as equation 8

$$\left|\frac{\ddot{\theta}_0}{\theta_i}\,\frac{1}{u^2}\right|=\left|\frac{1}{1-\dfrac{u^2}{k(\gamma)}+j[2\zeta_0]u}\right| \qquad (17)$$

To determine $k(\gamma)$ by the graphical method, it is necessary to plot a family of curves with constant values of $k(\gamma)$ as parameters. This is analytically expressed in equations 9 and 18

$$\left|\frac{\ddot{\theta}_0}{\theta_i}\,\frac{1}{u^2}\right|=\frac{1}{1-\dfrac{u^2}{k_n(\gamma)}+j[2\zeta_0]u} \qquad (18)$$

Fig. 6 is a plot of the magnitude of the right side of equation 18 versus u with $k(\gamma)$ a a parameter. The frequency plot of the left side of equation 18 is a straight line with a slope of −12 decibel per octave when u is plotted on logarithmic co-ordinates. This function is also plotted in Fig. 6 for one value of $\ddot{\theta}_0/\theta_i$. The intersection of this straight line with the family of constant $k(\gamma)$ curves determines the values of $k(\gamma)$ which satisfy equation 18 as u takes on different values, holding the ratio $\ddot{\theta}_0/\theta_i$ constant. This ratio can be thought of as a gain which determines the zero decibel crossover frequency of the linear functions. If this ratio is changed, a new set of values of $k(\gamma)$ which satisfies the equation is determined; consequently, $k(\gamma)$ is a function of ω and $\ddot{\theta}_0/\theta_i$. Since the closed-loop response of the system can be determined by substituting the computed values of $k(\gamma)$ into equation 3 at the specified frequency, the effect of both the magnitude and frequency of θ_i on the closed-loop response can be determined readily. If the product $k(\gamma)f[k(\gamma)]$ equals 1, as it does for the nonlinear element in Fig. 5, then the right side of equation 18 is the same as equation 3; consequently, both the closed-loop response and $k(\gamma)$ are determined at the same time. For nonlinear elements in which $k(\gamma)f[(k\gamma)]$ is not equal to 1, $k(\gamma)$ and the closed-loop response must be determined separately. An example of a nonlinear element in which the equivalent-gain function, $f[(k(\gamma)]$, is not equal to $1/k(\gamma)$ is shown in the Appendix.

A Comparison of the Analytical and Experimental Results

The log-modulus frequency response of the acceleration-limited system and the corresponding values of $k(\gamma)$ are derived in Fig. 6 for one value of the ratio $\ddot{\theta}_0/\theta_i$. The gain of the nonlinear element is 1 until $\gamma>\ddot{\theta}_0$; consequently, the closed-loop response will be on the $k(\gamma)=1$ curve until the function $(\ddot{\theta}_0/\theta_i)(1/u^2)$ intersects the $k(\gamma)=1$ curve. At this point $\gamma=\ddot{\theta}_0$, and this is the limit of linear operation. The frequency at the transition point will

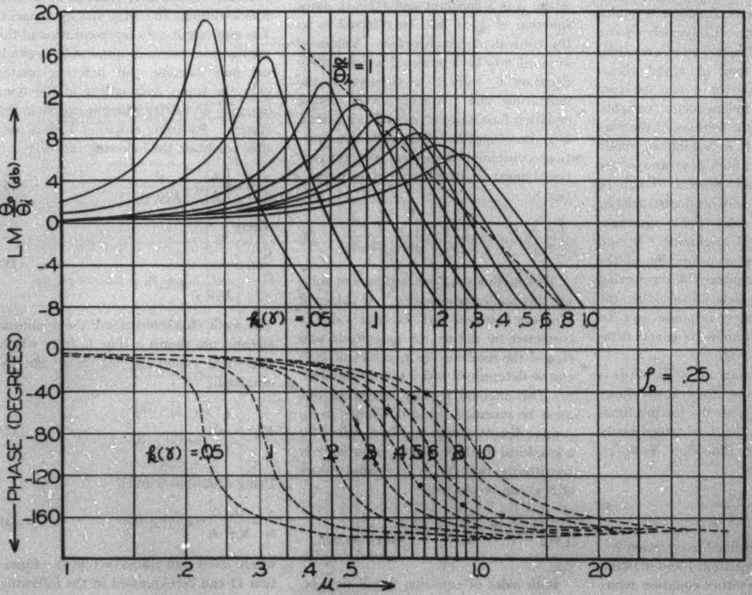


Fig. 6. Constant $k(\gamma)$ curves for function of equation 18

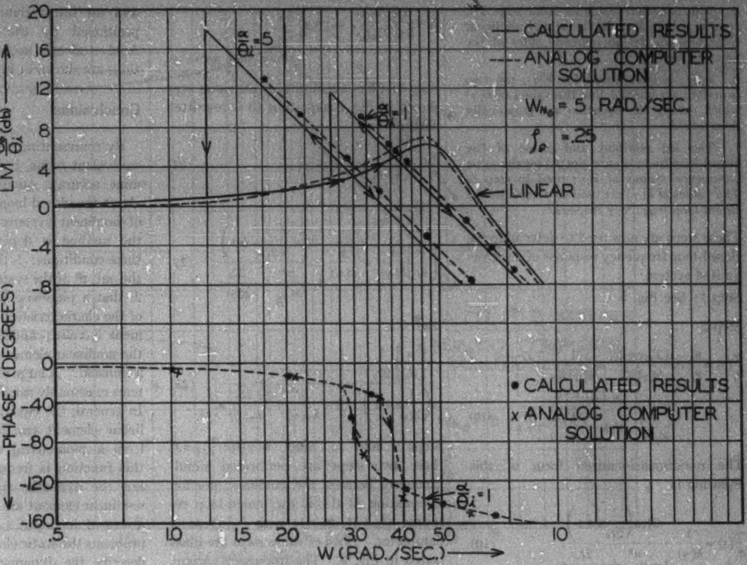


Fig. 7. Closed-loop frequency response of an acceleration-limited system

be referred to as U_T. Values of $k(\gamma)$ which satisfy the equation for $U>U_T$ are determined from intersections of the linear function and the $k(\gamma)$ curves. Since $k(\gamma)f[k(\gamma)]=1$ for this system, the frequency response of the system will also lie along the linear function locus which determines the values of $k(\gamma)$.

In Fig. 6 it can be seen that there are intersections between $(\ddot{\theta}_0/\theta_i)(1/U_2)$ and constant $k(\gamma)$ curves are frequencies lower than U_T; however, these values of $k(\gamma)$ are not realizable until nonlinear operation is established, i.e., when the acceleration is being limited. This phenomenon is commonly referred to as the response hysteresis or jump frequency, and it occurs because there are frequency regions in which $k(\gamma)$ is multivalued. The phase shift of the system can be determined from a family of phase curves computed from equation 14 with $k(\gamma)$ replaced by $k_n(\gamma)$. By using the values of $k(\gamma)$ determined from equation 18, the phase is determined by plotting points obtained from intersections of the constant $k(\gamma)$ curves and the specified values of U. These points are also seen in Fig. 6, for one value of $\ddot{\theta}_0/\theta_i$. Analytical and analogue-computer solutions are plotted in Fig. 7 for several different values of $\ddot{\theta}_0/\theta_i$ in order to determine the error introduced by approximating the nonlinear element by an equivalent gain.

It can be seen that there is general agreement between the approximate solution and the exact results obtained from the analogue computer. During the experimental work, it was not possible to obtain as many points in the nonlinear region for $U<U_T$ as indicated by the analysis. The reason for this is not clearly understood at present. Considerable difficulty was encountered in obtaining these points with the computer as this appears to be a semiunstable region. It can be seen that the closed-loop response in the nonlinear region is determined by the open-loop response of the elements to the right of the nonlinearity in Fig. 4; consequently, it is not surprising that the nonlinear response shows an attenuation of 12 decibels per octave, since the output of the limiter is constant during nonlinear operation.

Outline of the Procedure

To formulate a general procedure for determining the closed-loop frequency response, the following steps are applied to an error-limited system:

1. The system equations are expressed, including the nonlinear element, as transfer functions, and the block diagram is drawn.
2. The closed-loop transfer function is determined, relating the input to the nonlinear element γ to the system input θ_i.
3. An analytical expression is determined relating the input γ and the equivalent gain, $k(\gamma)$, of the nonlinear element with the use of only the static characteristics. Both sides of the equation are divided by θ_i.
4. With the use of the magnitude of the equations obtained from steps 2 and 3, these equations are solved simultaneously. γ is eliminated. The result will be a nonlinear equation with $k(\gamma)$ as the only unknown.

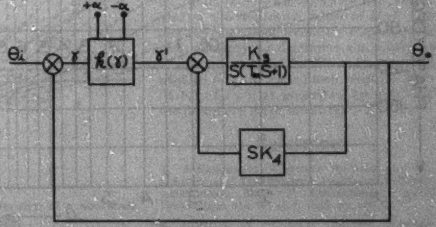


Fig. 8. An error-limited system

5. The equation obtained from step 4 is rearranged so that the linear and nonlinear functions will be on opposite sides of the equality sign.

6. $k_n(\gamma)$ is substituted for $k(\gamma)$, the two functions are plotted, and the values of $k(\gamma)$ which satisfy the equation are graphically determined.

7. The log modulus and phase of the closed-loop function are plotted for the same parametric values of $k(\gamma)$ used in step 6. The values of $k(\gamma)$ are used to determine the closed-loop frequency response.

These steps are now used to determine the closed-loop frequency response of an error-limited system.

Step 1: See Fig. 8

Step 2

$$\frac{\gamma}{\theta_i}(s)=\frac{\theta_0}{\theta_i}(s)\frac{s(T_ms+1)}{k(\gamma)K_3}=\frac{1}{k(\gamma)K_3}\times \frac{s(T_ms+1)}{\dfrac{T_m}{K_3k(\gamma)}s^2+\dfrac{1+K_3K_4}{K_3k(\gamma)}i+1} \qquad (19)$$

The nondimensionalized form of this equation is

$$\frac{\gamma}{\theta_i}(s)=\frac{1}{k(\gamma)}\,\frac{2\zeta_0u\left(\dfrac{1}{2\zeta_0}u+1\right)}{1-\dfrac{u^2}{k(\gamma)}+j\,\dfrac{2\zeta_0}{k(\gamma)}\,u} \qquad (20)$$

where

$$\zeta_0=\frac{1+K_3K_4}{2\sqrt{T_mK_3}},\quad \omega_{n_0}=\sqrt{\frac{K_3}{T_m}},$$

$$s=j\omega,\ \text{and}\ u=\omega/\omega_{n_0}$$

Step 3: Since the limiter is the same type shown in Fig. 5, equation 13 is rewritten with $\alpha=\ddot{\theta}_0$.

$$\frac{\gamma}{\theta_i}=\frac{\alpha}{k(\gamma)}\,\frac{1}{\theta_i}\ \text{for}\ |\gamma|>|\alpha| \qquad (21)$$

Step 4

$$\left|\left[\frac{\alpha}{\theta_i}\right]\frac{1}{k(\gamma)}\right|=\left|\frac{1}{k(\gamma)}\cdot\frac{2\zeta_0u\left(\dfrac{1}{2\zeta_0}u+1\right)}{1-\dfrac{u^2}{k(\gamma)}+j\,\dfrac{2\zeta_0}{k(\gamma)}u}\right| \qquad (22)$$

Step 5

$$\left|\frac{\left[\dfrac{\alpha}{\theta_i}\right]\dfrac{1}{2\zeta_0}}{u\left(\dfrac{1}{2\zeta_0}u+1\right)}\right|=\left|\frac{1}{1-\dfrac{u^2}{k_n(\gamma)}+j\,\dfrac{2\zeta_0}{k_n(\gamma)}u}\right| \qquad (23)$$

Steps 6 and 7: Since $k(\gamma)f[k(\gamma)]=1$, these two steps are performed simultaneously because the nonlinear portion of equation 23 also is the closed-loop response $(\theta_0/\theta_i)[s_1k_n(\gamma)]$ with $k_n(\gamma)$ as a parameter. Both of these steps are illustrated in Fig. 9. The frequency-variant left portion of equation 23 can be determined from a template of a first-order lag and an integration. This template is positioned by the gain $(\alpha/\theta_i)(1/2\zeta_0)$. Analytical and analogue computer solutions are shown in Fig. 10.

Conclusions

By representing nonlinear elements by equivalent gains, it is possible to determine accurate quantitative information about the closed-loop frequency response of nonlinear systems. The accuracy of the method is dependent primarily on three conditions: 1. that the harmonics at the output of the system can be neglected; 2. that a piecewise linear approximation of the characteristics of the nonlinear element is valid; and 3. that the input to the nonlinear element is representable by a sinusoid. A large class of nonlinear systems reasonably satisfies these conditions. In general, the equivalent gain of a nonlinear element can be computed easily from a geometrical construction. Also, this function is frequency invariant unless the dynamic characteristics of the nonlinear element are a function of derivatives of its input. For most practical problems the static characteristics equally describe the dynamic operation; consequently the equivalent-gain function remains frequency invariant even though

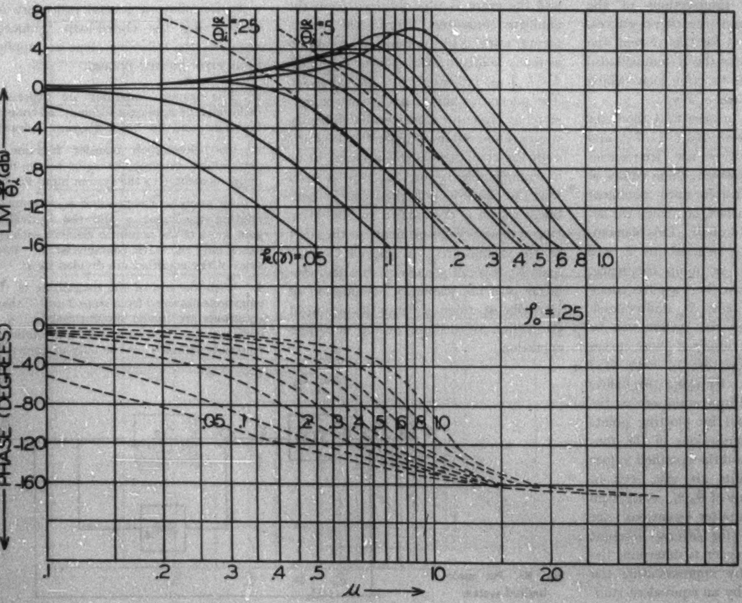


Fig. 9. Constant $k(\gamma)$ curves for the function of equation 23

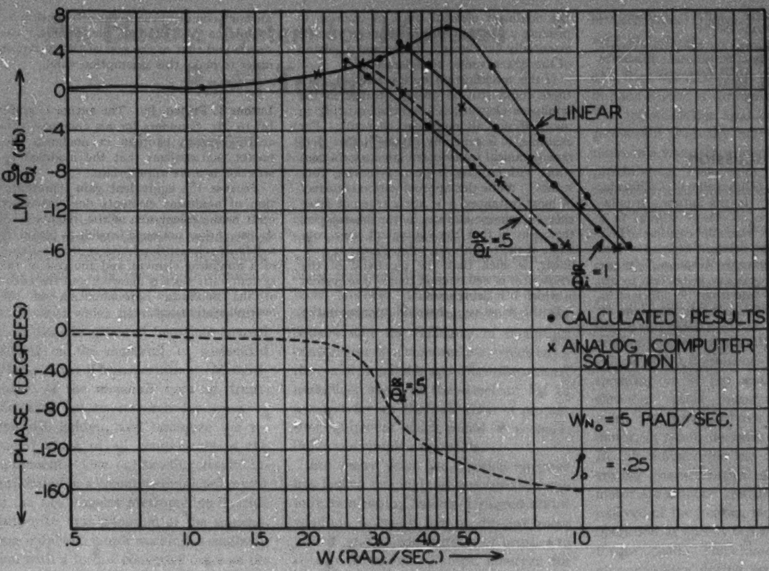


Fig. 10. Closed-loop frequency response of an error-limited system

the nonlinear element may be located in a minor loop containing energy-storage elements. Although this method was applied only to second-order systems, it is equally applicable to higher order systems as long as only one nonlinear element is significant. Another advantage provided by the equivalent-gain approximation is that the effect of changes in the closed-loop frequency response, caused by a change in the magnitude of θ_i, can be easily determined by positioning the template of the left side of equation 23 in accordance with various values of θ_i.

Since this method of analysis has been found to be reasonably accurate, and also since the method does not inolve procedures which are unfamiliar to feedback-systems engineers, it is felt that it may prove to be suitable for general engineering use. It should also prove to be of value in realizing the full potential of the sinusoidal response method of analyzing nonlinear feedback-control systems.

Appendix

An analytical expression for the equivalent gain of the limiter used in the text was derived in equation 12. A more general expression, however, was indicated in equation 1

$$k(\gamma)=\frac{\gamma'}{\gamma}$$

which can be rewritten as

$$\gamma=\frac{\gamma}{k(\gamma)}=\alpha f[k(\gamma)]$$

The term $f[k(\gamma)]$ has been defined as the equivalent-gain function, which equals $1/k(\gamma)$ for the type of limiter used in the text. This is a special case of a nonlinear function, as the product $k(\gamma)f[k(\gamma)]$ is not equal to 1 in the general case; consequently, it will be shown how nonunity values of this product affect the solution of the nonlinear equation. The static characteristics of a more general nonlinear element are shown in Fig. 11, and the equivalent-gain function for this element will now be determined. The following relationship can be geometrically derived.

$$\gamma'=\gamma[k(\gamma)]=\alpha+K_1(\gamma-\alpha) \quad \text{for } |\gamma|>|\alpha| \qquad (24)$$

$$\gamma[k(\gamma)-K_1]=\alpha(1-K_1) \qquad (25)$$

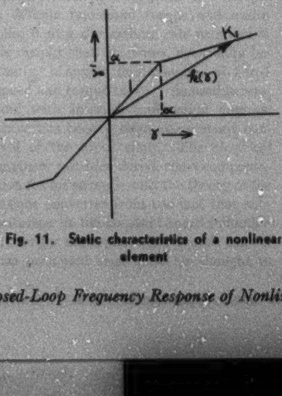


Fig. 11. Static characteristics of a nonlinear element

consequently

$$\gamma=\frac{\alpha(1-K_1)}{k(\gamma)-K_1}=\alpha f[k(\gamma)] \qquad (26)$$

In this case

$$f[k(\gamma)]=\frac{1-K_1}{k(\gamma)-K_1}$$

If this value of $f[k(\gamma)]$ is substituted in equation 9, the equation becomes

$$\left|\left[\frac{\alpha}{\theta_i}\right]K_3G_2(s)\right|=\left|\frac{\theta_o}{\theta_i}[s,k_n(\gamma)]\left\{\frac{k(\gamma)-K_1}{k(\gamma)[1-K_1]}\right\}\right| \qquad (27)$$

In this equation, as in equation 9, $k(\gamma)$ is the only unknown when the magnitude and frequency of θ_i are specified; consequently, the values of $k(\gamma)$ which satisfy the equation can be determined by the method which has already been outlined. The left side of the equation will remain unchanged, and the only difference on the right side is that the curves, determined by the function, $(\theta_o/\theta_i)[s,k_n(\gamma)]$, will be displaced from each other by a gain, $[k(\gamma)-K_1]/[k(\gamma)(1-K_1)]$, evaluated at the respective values of $k(\gamma)$. To plot the closed-loop frequency response, it is first necessary to determine the values of $k(\gamma)$ from equation 27. The values are then used to plot the closed-loop response from $(\theta_o/\theta_i)[s,k_n(\gamma)]$, as prescribed by step 7 of the procedure listed in the text.

References

1. A Frequency Response Method for Analyzing and Synthesizing Contractor Servomechanisms, Ralph J. Kochenburger. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

2. Limiting in Feedback Control Systems,

Ralph J. Kochenburger. *AIEE Transactions*, vol. 72, pt. II, July 1953, pp. 180–94.

3. SINUSOIDAL ANALYSIS OF FEEDBACK-CONTROL SYSTEMS CONTAINING NONLINEAR ELEMENTS. **E. Calvin Johnson.** *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 169–81.

◆

Discussion

R. M. Saunders (University of California, Berkeley, Calif.): In the author's stipulation that there be no energy storage in a nonlinear element, which appears in the first paragraph of the section entitled "The Closed-Loop Frequency Response," it appears that an unnecessary restriction to his work is made. Would it not be sufficient to state that the equivalent time constant of the nonlinear element must be such as to produce a passband for the element which is essentially flat through the normal passband of the system under consideration?

If this modification were permitted, then there could still be energy storage in the nonlinear element, a condition existing in many if not all nonlinear elements. For example, it is generally supposed that electromechanical transducers have a gain constant without any frequency-varying constants. While this approximation is correct in most instances, it is not a true one since there is energy storage in the transducer; the point is that the equivalent time constant of the transducer, either linear or nonlinear, is such that the passband of the transducer is well beyond that of the system in which it is being used.

In this same vein, I would appreciate the author's comments as to the relationship between the pass-band of the nonlinear elements and that of the rest of the system in order to make this assumption valid.

Luther T. Prince, Jr.: The author is grateful to Mr. Saunders for his comment concerning energy storage in nonlinear elements, and realizes that the modified restriction is more appropriate.

Because the equivalent gain representation of nonlinear elements does introduce error in the calculation of the frequency response, it does not seem feasible to postulate a definite relationship between the passband of a nonlinear element and the rest of the system until more is known about the effect of the nonlinear approximation on the over-all system error.

An Electric System for Monorail Rapid Transit

E. H. ANSON
MEMBER AIEE

R. L. KIMBALL
ASSOCIATE MEMBER AIEE

SURFACE traffic congestion and the high cost of subway construction have caused renewed interest in elevated mass transit. The necessity to eliminate noise and to avoid shutting off light and air, both of which have been characteristic of the old elevated lines has focused interest on the monorail, or suspended railway, form of transportation. Modernization of the monorail type of transportation led to research for ways to simplify, lighten, and minimize the required space of all components of the equipment. One outstanding result, incorporated in a recently completed report to the Los Angeles Metropolitan Transit Authority, was adoption of the squirrel-cage induction motor working in combination with a torque converter drive as the propulsion equipment, in place of the d-c series motor and its involved rheostatic voltage control. The resulting performance characteristics are the equal of the most recent series motor equipment proposed for comparable service. Besides weight and space advantages on the rolling stock, a substantial reduction in power supply apparatus costs results from the direct use of usual 3-phase 60-cycle electric energy without the expense and the complication of special conversion to direct current.

Propulsion of monorail suspended-railway passenger vehicles by means of a number of different media has been suggested. The schemes mentioned run the gamut from a ram-jet, through motor-driven or engine-driven propellers, to electric motor drive. The monorail's objective of an unusually quiet form of transportation at high speed, which in turn implies high horsepower capacity, points definitely to the choice of electric motor drive. Jet or propeller propulsion would be entirely too noisy for serious consideration in urban areas, while even internal combustion engine drives, such as are used on the bus, would be objectionably noisy to many persons, especially when trains of cars are operated. Further, jet and engine drives both emit objectionable fumes from combustion which can aggravate smoke-fog conditions.

The d-c series motor and control equipment, developed for light-weight streetcar propulsion, is available and suitable for monorail application, and initially was considered to be the natural solution. Mass transportation systems everywhere have, however, demonstrated the necessity of reducing maintenance costs. This feature and the difficulty of mounting control gear on a monorail car, which lacks the readily available and otherwise unused underfloor space of the streetcar, led to a search for an alternate form of propulsion equipment which would be as small, light, and simple as possible.

For many years the traction engineer has looked longingly at the squirrel-cage induction motor having a rugged stator winding, a heavy bar or die-cast rotor circuit, and no commutator or brushes to require attention. Except, however, for specialized locomotive applications, the problems of acceleration, speed control, and current collection have denied the adoption of the induction motor for rail-borne traction.

In recent years the automotive field has produced a solution to the first of these objections, in the form of the torque converter drive. Monorail itself, by its nature and form, removes the other two difficulties.

Although the fine points of the theory and design of a torque converter may present somewhat of a mystery to electrical engineers, the characteristics of its performance are generally familiar from experience with automotive and modern rail car operation. In a manner of oversimplification the torque converter can be regarded as a variable gear ratio which at times of maximum input-to-output speed differential across the unit accepts a definite maximum torque and multiplies it at a diminishing rate as speed of its output shaft increases, eventually becoming a simple fluid coupling at 1-to-1 speed and torque ratios. As used heretofore with internal combustion engines, there has been present some speed control of the input shaft. The electrical engineer can here derive some compensation for his struggle with the theory of the torque converter from the fact that substitution of the constant-speed induction motor for the engine with throttle control has presented some food for thought to his mechanical confrere. The outcome is very interesting.

Assume that the vehicle is at standstill and that the 3,600-rpm squirrel-cage induction motor is thrown across the line. The initial torque load on the motor is zero and it gradually increases to its maximum only when the motor has reached approximately 2,880 rpm. Further, this maximum load imposed on the motor can be considerably below its breakdown torque. As a result, the motor is running well within its desirable operating-speed range in a very short time, and maximum driving torque, in the order of 3.5 times the motor torque, is applied to the drive shaft. The converter output driving torque is gradually reduced as the vehicle speed approaches a value corresponding to 80 per cent of motor synchronous speed, presumably about 50 miles per hour. Above this speed the motor, in effect, operates in direct drive. In this manner the less desirable, accelerating characteristics of the induction motor are avoided and the motor itself is relieved of the heating incident to loaded operation at speeds below its breakdown torque speed. The elimination of resistance notching during acceleration is an incidental but important feature in the reduction of first cost and maintenance expense.

Because the monorail operates on a private right of way free of impeding cross traffic, its need of low-speed running notches is reduced, although for use in yard switching and a few speed-restricted locations on the main line a lower acceleration and a half-speed running notch are desirable. The half-speed is readily obtainable by pole reconnection of the motor to provide a synchronous speed one-half of that normally used. When a torque converter is operated at lower speeds its torque multiplication is reduced. This reduced multiplication accompanying the half motor speed reduces the normally high rate of acceleration to a value suitable for making up trains, maneuvering in yards, and "closing-up" movements on the main line.

Because backward rotation for reverse moves is not available from internal combustion engines, application of the torque converter to engines usually also includes a reversing gear. The induction motor can, however, be readily reversed.

Paper 54-251, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 23, 1954; made available for printing May 17, 1954.

E. H. ANSON and R. L. KIMBALL are with Gibbs and Hill, Inc., New York, N. Y.

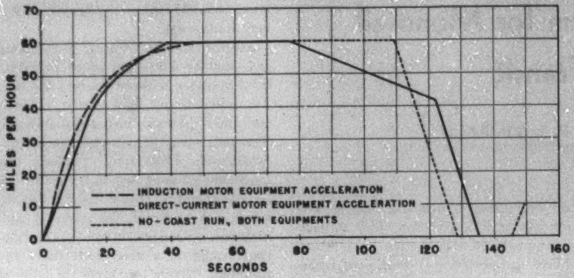


Fig. 1. Comparative speed-time curves

Length of run 1.76 miles
Scheduled speed 43.5 miles per hour; speed margin 5 per cent

Careful examination by the Allison Division of the General Motors Corporation revealed its torque converter could be adapted for reverse drive but without torque multiplication. Thus, such few reverse moves as are required and dynamic "holding" braking if it is desired, can be obtained by simple motor reconnection.

The supporting structure of the monorail provides ample facility for mounting a dual contact system in an inconspicuous location. The running rail serves as the third phase-conductor supplying 3-phase 60-cycle energy, probably at 2,300 volts, to the trucks. The necessary control apparatus, consisting of a main interrupting switch and a pole-changing cam group for speed-changing and reversing, can be mounted on the truck, thus avoiding the necessity of carrying propulsion power circuits down into the car body.

Fig. 1 illustrates the performance of equipment using the induction motor drive, included in the report to the Los Angeles Metropolitan Transit Authority, and a closely comparable d-c series-motor-driven equipment recently proposed for another rapid-transit operation. The latter equipment might even be considered to be overmotored because high-capacity motors were proposed to obtain fast acceleration to a top operating speed of 60 miles per hour. It can readily be seen that, in spite of this condition, the induction motor and torque converter drive delivers substantially the same car performance. Both equipments are capable of operating a schedule of 43.5 miles per hour with a 1.76-mile average length of run and with a speed margin of 5 per cent. It is therefore evident that adoption of the induction motor drive has not sacrificed anything in equipment performance.

From the viewpoint of cost, and on the basis of preliminary design figures, it is not expected that there will be great differences in propulsion equipment or power distribution costs between the two systems. The lesser space requirement of the induction motor propulsion apparatus is reflected in the lower over-all height of equipment, and therefore in the lesser cost of supporting structures. A greater difference lies in the cost of power supply facilities. The induction motor operates on energy of standard commercial frequency and voltage, while the conventional series motor requires conversion of energy supplied from commercial a-c sources to direct current.

To supply the induction motor drive, power will be taken from the utility supply through simple conventional unit-type transformer substations placed adjacent to passenger stations where the maximum demands of acceleration occur. Several of these transformer stations, to the extent permitted by characteristics of the supplying system, would function together to supply the contact system as a line network to which all transformers would deliver power and share loads to the extent permitted by the transmission ability of the contact system. This feature is especially important in the event of the outage of one of the transformers or its source. The cost of transformer stations is estimated as slightly below $20 per installed kilovolt-ampere.

A d-c supply system would undoubtedly employ rectifier substations the transformer of which is more expensive and which has in addition a mercury rectifier or other type of rectifier, with its control and protective equipment, all adding up to $80 or $100 per kilowatt of installed capacity. Although the d-c installed kilowatts of conversion capacity would be approximately 80 per cent of the a-c installed kilovolt amperes, the difference is not sufficient to overcome the handicap of higher unit cost.

◆

Discussion

W. H. T. Holden (Consultant, Pasadena, Calif.): Full subway or overstreet elevated lines are not the only possible constructions with which to solve traffic congestion, as might be inferred from the first sentence of the paper. That this is not necessarily the case has been brought out by Vouch.[1] Other types of construction are known and in use, for example, the Brighton Beach and Sea Beach lines of the New York City system which occupy open cut.

However, the problem of electrical propulsion of a vehicle is somewhat independent of the geometry of the support of the vehicle. The system here proposed may be regarded as having certain inherent disadvantages. The principal disadvantage is the irreversible nature of the drive unless gear reverse is added. While many street railways have used single-ended equipment, no rapid transit line has been designed on this basis in the United States or Canada. The need for loops to turn trains and the inability to secure emergency turn-back would be a serious problem. In the design of monorail discussed in the report to the Los Angeles Metropolitan Transit Authority this single-ended limitation is inherent in the suspension. While many newer car designs are single-ended to some degree, a double-ended operating unit is secured by semipermanent coupling of these cars in pairs. Such units are in use on the IRT division of the New York system in Chicago, and in Toronto. The propulsion system here described is not applicable to this type of double unit.

The squirrel-cage induction motor proposed is understood to have a low-resistance rotor. Except for a limited range of speeds, such motors cannot furnish dynamic braking by d-c excitation of the stator, and in this case the energy is liberated in the rotor. As dynamic braking at high torque is regarded as highly desirable in rapid-transit applications, the inability of this equipment to provide high-torque dynamic braking must be regarded as a serious disadvantage.

It may be suggested that the complexity of rheostatic control is not a serious item. It is also suggested that an appreciable amount of switching is required to shift from bipolar to 4-pole connections. As a maintenance item, it is believed that the torque converter will be comparable to the brushes of a wound-rotor motor.

The use of the torque converter has resulted in an induction motor drive that closely duplicates the series motor speed-time curve. One of the limitations of d-c railway drives using series motors is the lack of accelerating capability at speeds above about 50 per cent of balancing speed. At the cost of a substantial increase in power consumption, the wound rotor induction motor allows full-rate acceleration to about 90 per cent of running speed, has a reduced running speed possible by cascade connection, and can regenerate if no gaps exist in the contact line at least from full speed down to about 60 per cent of full speed, and can furnish high torque dynamic braking down to very low speed by d-c stator excitation. Such a system has the advantages of direct utilization of 60-cycle power, as does the system proposed by the authors.

Reference

1. High-Speed Rapid-Transit Equipment, S. J. Vouch. *AIEE Transactions*, vol. 73, pt. II, Nov. 1954, pp. 267–70.

Some Fundamentals of Equipment-Grounding Circuit Design

R. H. KAUFMANN
FELLOW AIEE

Synopsis: An effective equipment-grounding system should, under conditions of maximum ground-fault current flow, accomplish the following objectives: 1. maintain a low potential difference, perhaps 50 volts maximum, between machine frames, equipment enclosures, conductor enclosures, building metallic structure, and metallic components contained therein to avoid electric shock hazard and unwanted circulating current, and 2. incorporate adequate conductance to carry this maximum ground short-circuit current without thermal distress and the attendant fire hazard. There is good reason to believe that current flow in the equipment-grounding system in a-c power systems will not stray far from the power cable over which the outgoing current flows. It follows that the installation of conductive material in an equipment ground system unless properly located can be ineffective and wasteful, and can create a false sense of security. This paper presents the results of a special series of full-scale tests dealing with this specific problem and a general analysis of the circuit behavior.

A MUCH better understanding of protective grounding systems is necessary to ensure freedom from hazard to life and property. The objective of this paper is to present the significant factors which control the behavior of protective grounding circuits in a-c industrial power distribution circuits during short-circuit conditions. It is hoped that this presentation may emphasize that improper application of material in equipment-grounding systems creates the wastefulness and unwarranted sense of security just mentioned.

Under a short-circuit to ground condition in an a-c distribution circuit, grounded or ungrounded systems alike, inductive reactance will exert a powerful influence in directing the return current to flow in a path closely paralleling the outgoing power conductor. The enclosing conduit or metallic raceway would constitute such a path. To attempt to relieve the conductor enclosure by installation of an external conductor is quite ineffective. Connections to near-by building structural members are equally ineffective. Only through the installation of an internal grounding conductor can the current which must be carried by the enclosure be noticeably reduced. Joints in conduits and raceways require special consideration to avoid a shower of sparks and attendant fire hazard during the short circuit's duration. Any thought of keeping the conduit or raceway insulated from ground except at one point seems totally impractical. Furthermore, such practice or a contemplated omission of the metallic enclosure may well lead to large induced voltages in near-by metallic structures which may appear either as dangerous shock hazards or unwanted circulating currents. Ungrounded distribution systems require equally careful treatment. Very often a second ground fault occurs before the previous one has been located and corrected. The problems discussed in this paragraph then appear simultaneously on each of the two circuits.

Test Procedure

A special installation of $2^{1}/_{2}$-inch conventional heavy-wall steel conduit and 4/0 copper cable conductors was made for this investigation. It was installed in a building previously used for short-circuit testing because this building contained a heavy steel column construction and all columns were tied to an extensive grounding mat composed of 250,000-circular-mil copper cables. The test installation is illustrated in Fig. 1. The conduit was supported on insulators throughout the 100-foot length. The conduit run was about 5 feet from a line of building columns. The external 4/0 cable was spaced about 1 foot from the conduit on the side opposite the building columns.

The particular arrangement of components makes possible the measurement of all currents and voltages at one location. This is especially valuable in that it avoids the running of lengthy voltage-measuring leads. The setup is intended to simulate an electric feeder circuit with power source at the left end and various simulated fault conditions at the right end. In all cases here reported the test current was caused to flow over the *A* cable to the far end. A variety of different possible return paths were examined, controlled by the connections made at the left end.

One series of tests was made at low current, 200 and 350 amperes, using an a-c welding transformer as a source of 60-cycle power. At these low-current magnitudes, the current flow could be maintained for extended periods. Voltage measurements were made with high quality indicating voltmeters. Current measurement was made with a clip-on ammeter.

A second series of tests was made at high current, around 10,000 amperes, using a 450-kva 3-phase 60-cycle transformer with a 600-volt secondary as a source of power. Switching was done at primary voltage, 13,800 volts, and an *IAC* induction relay was used to control the duration of current flow to an interval of about 1/4 second. An oscillograph was used for all measurement of current and voltage.

Paper 54-244, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 23, 1954; made available for printing April 27, 1954.

R. H. Kaufmann is with the General Electric Company, Schenectady, N. Y.

The generous assistance of J. M. Schmidt and his staff in Facilities Engineering of the General Electric Company was invaluable to the test portion of this investigation. To him goes the credit for making the test installation, arranging for the loan of power supply transformers, together with switching and relaying equipment, and, in addition, authorizing the short-circuit tests from the 13,800-volt feeder lines serving the Schenectady plant. To L. J. Carpenter of Industrial Power Engineering goes credit for assistance in conducting the tests and the analysis of results.

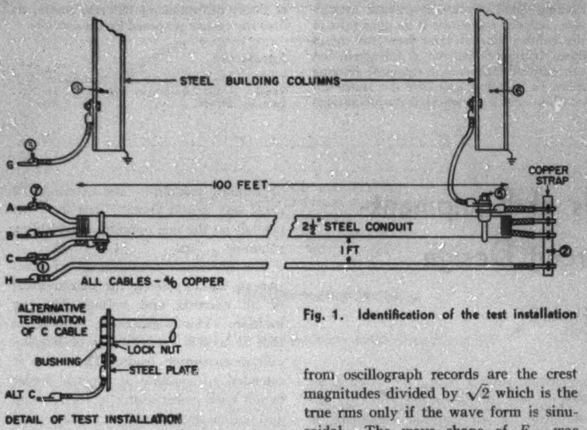


Fig. 1. Identification of the test installation

Results

The magnitudes of voltage and current obtained in the entire series of 60-cycle a-c tests are presented in organized form in Table I. The test number is an arbitrary assignment for reference purposes. The next two columns identify the connections used and indicate the possible paths of current flow. Next the current values are tabulated: first, the total input current into the *A* conductor and, next, the return current in the conduit and its percentage of the total; and then the magnitudes of current in the other possible paths are given.

In all cases the tabulated values taken from oscillograph records are the crest magnitudes divided by $\sqrt{2}$ which is the true rms only if the wave form is sinusoidal. The wave shape of E_{CG} was commonly quite different from a sine wave, but in the extreme cases the magnitude also was small. See Fig. 3 pertaining to test B8.

Measurements of d-c resistance were made of the several return paths comprising the equipment-grounding system. See Table II. A controlled magnitude of direct current was caused to flow through the circuit and the d-c potential drop across the desired section was measured with an indicating millivoltmeter. The points of measurement refer to those indicated in Fig. 1.

No further analysis is needed to show conclusively that only by the use of an internal grounding conductor can any sizable fraction of the return current be diverted from the conduit or raceway. In spite of the extremely low resistance of the building structural frame, it was ineffective in reducing the magnitude of return current in the conduit. See tests A6, A7, B6, and B7. Further analysis is desirable to establish a better understanding of the nature of the circuits involved.

Some interesting secondary effects were observed in the course of the tests. The first high-current test produced a shower of sparks from about half the couplings in the conduit run. From one came a blowtorch stream of sparks which burned out many of the threads. Several small fires set in near-by combustible material would have been serious if not promptly extinguished. The conduit run had been installed by a crew regularly engaged in such work and they gave assurance that the joints had been pulled up to normal tightness and perhaps even a little more. A short 4/0 copper jumper was bridged around this joint but, even so, some sparks continued to be expelled from this coupling on subsequent tests. Other couplings threw no more sparks during subsequent tests. Apparently small tack welds had occurred on the first test.

In one high-current test the conduit termination was altered to simulate a connection to a steel cabinet or junction box. See bottom of Fig. 1. The bushing was applied finger-tight. On test with about 11,000 amperes flowing for about 1/4 second, a fan-shaped shower of sparks occurred parallel to the plate. In the process, a weld resulted and the parts were separated only with con-

Table I. Measured Electrical Quantities

	Current Flow		Current Magnitudes						Voltage Magnitudes					
				I_C										
Test No.	Out On	Return On	I_A Total	Amperes	Per Cent of Total	I_B	I_H	I_G	E_{AC}	E_{CG}	E_{AB}	E_{AH}	E_{AG}	E_{GB}
						Low-Current Tests								
A1	A	B	350	0		350	0	0	2.47		4.85			
A2	A	C	350	350	100	0	0	0	15.9	0.45	2.5			
A3	A	C	200	200	100	0	0	0	9.05	0.15	1.51			
A4	A	CH	350	340	97	0	12	0	16.0	0.05	2.55			
A5	A	CH	200	190	95	0	8	0	9.13	nil	1.55			
A6	A	CG	350	340	97	0	0	12	14.6		2.54	14.4		
A7	A	CG	200	180	90	0	0	8	9.5		1.50	9.4		
A8	A	CB	350	62	18	290	0	0	4.55	nil	4.55			
A9	A	CB	200	40	20	150	0	0	2.68	nil	2.68			
A10	A	GH	350	0	0	0	160	160	14.0	12.5	2.5	26.4	26.4	24.6
A11	A	GH	200	0	0	0	98	98	9.2	8.1	1.5	17.1	17.1	15.1
						High-Current Tests								
B2	A	C	11,200	11,200	100	0	0	0	168	36*				
B3	A	C	11,070	11,070	100	0	0	0	173	38*				
B4	A	CH	11,070	11,200	101	0	1,140	0	173	18*				
B5	A	CH	11,080	11,090	100	0	1,220	0	173	17*				
B6	A	CG	10,830	10,770	99	0	0	1,080	168		71			
B7	A	CG	10,910	10,780	99	0	0	1,145	173	9*				
B8	A	CB	11,620	5,810	50	5,660	0	0		27*			155	
B9	A	CB	11,380	6,070	53	5,620	0	0	146	25*				
B10	A	GH	8,710	0	0	0	4,300	4,500		146			268	

* Distorted wave shape. Tabulated values are crest/$\sqrt{2}$.

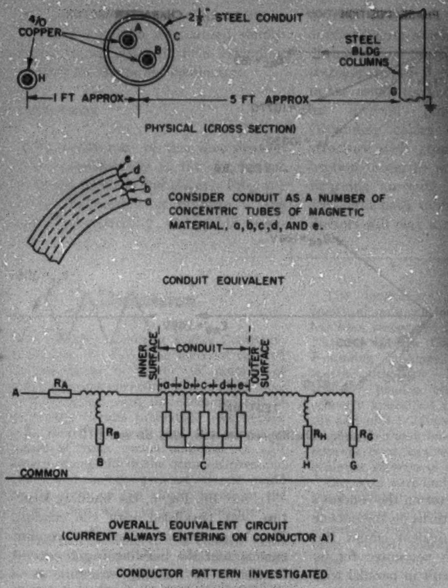

4/0 COPPER
$2\frac{1}{2}$" STEEL CONDUIT
A
B
C
H
STEEL BLDG COLUMNS
G
1 FT APPROX
5 FT APPROX
PHYSICAL (CROSS SECTION)
a
b
c
d
e
CONSIDER CONDUIT AS A NUMBER OF CONCENTRIC TUBES OF MAGNETIC MATERIAL, a, b, c, d, AND e.
CONDUIT EQUIVALENT
INNER SURFACE
CONDUIT
OUTER SURFACE
A
R_A
R_B
B
a
b
c
d
e
C
R_H
H
R_G
G
COMMON
OVERALL EQUIVALENT CIRCUIT (CURRENT ALWAYS ENTERING ON CONDUCTOR A)
CONDUCTOR PATTERN INVESTIGATED

Fig. 2. Test circuit conductor geometry and equivalent electric circuit

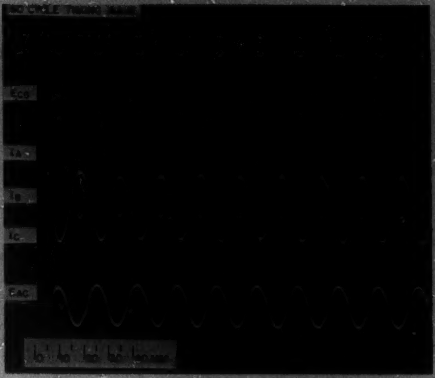


Fig. 3. Oscillogram of electrical quantities in test B8

E_{CO} conduit to ground volts, 1.26 volts per millimeter
I_A line current, 785 Angstrom units per millimeter
I_B ground cable current, 623 Angstrom units per millimeter
I_C conduit current, 247 Angstrom units per millimeter
E_{AC} input volts, 25.8 volts per millimeter

siderable difficulty, with the use of wrenches and a hammer. This suggested that a repeat shot would have produced no disturbance.

During high-current test B10 (conduit circuit open) a shower of sparks was observed at an intermediate building column. Careful inspection disclosed that the origin was at a spot at which a water pipe passed through an opening cut in the web of the steel beam involved. The pipe had been loosely in contact with the edges of the hole. Here is evidence of the objectionable effects of forcing the short-circuit current to seek return paths remote from the outgoing conductor. The large spacing between outgoing and returning current creates a powerful magnetic field which extends far out in space around the current-carrying conductors.

Circuit Analysis

The test installation as identified in Fig. 1 allows a study of a wide variety of equipment-grounding arrangements. The *A* conductor in all cases is used to represent the power conductor which has faulted to ground. One side of the test power was connected directly to the *A* conductor in all tests. The *B* conductor can be connected to act as an internal grounding conductor. The *H* conductor can represent an external grounding cable run parallel to the power cables with about 1 foot separation.

In addition to a resistance value associated with every conductor element is an inductive reactance value of a-c impedance. The reactance will increase as the spacing increases. Thus, for conditions in which the outgoing current will in all cases be carried by the *A* conductor, the reactances of the various elements will increase as their spacing from the *A* conductor increases.

The upper sketch in Fig. 2 defines the spacing characteristics of the circuit elements being studied. The steel conduit represents an annular area of high magnetic permeability. Until saturation occurs the magnetic flux produced in the steel pipe may be 500 or more times what it would have been in nonmagnetic material. The conduit may be considered as a series of thin-wall magnetic tubes, one within the other, such as that indicated in the center sketch in Fig. 2. The resulting equivalent circuit of the system under investigation takes the form indicated at the bottom of Fig. 2. The reactance of the circuit including the *B* conductor will be the lowest. Next will be the innermost tube of the conduit, followed by the others in successive order until finally the outer tube is reached. The inductance of these tubular elements of the steel conduit assumes unusual importance because of the high magnetic permeability. Next in spacing is the external grounding conductor (*H* cable) and, last, the structural members of the building frame and their interconnecting grounding cables buried below floor level. Any given circuit element is brought into use by connecting its terminal to the common bus.

Table II. Measured D-C Resistances

Element	Millivolt Drop Between	Ohms
100-foot length of 4/0 conductor	2–7	0.00515
100-foot length of $2^{1}/_{2}$-inch steel conduit	1–5	0.00475
Steel building frame	4–6	0.00011
	3–5	0.0006

The effectiveness of the conduit in confining the flow of fault current within the conduit can be visualized with the aid of this equivalent circuit. Consider only the conduit (*C* terminal) connected to the common bus. This forces all the current to return on the conduit alone. The equivalent circuit shows that the current by returning on the inner surface of the conduit will avoid the reactance of the other annular sections. Furthermore, the conduit reactance will impede

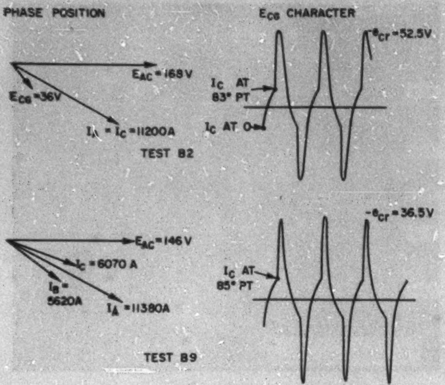


Fig. 4. Oscillogram analysis, tests B2 and B9

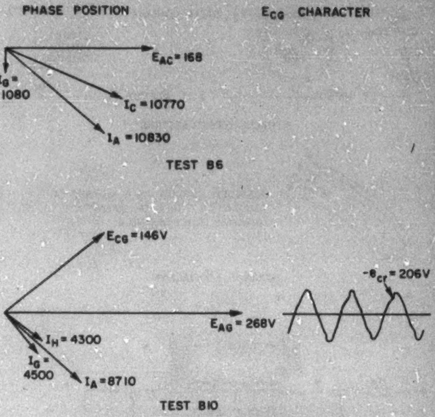


Fig. 5. Oscillogram analysis, tests B6 and B10

the diversion of current to the conduit outer surface. The voltage drop in a circuit path external to the conduit should be expected to be low until magnetic saturation occurs. Observe this effect in tests A2 and A3. The voltage E_{CG} is a voltage drop measured around a circuit external to the conduit. The value of 0.45 volt with 200 amperes flowing is less than the product of the current magnitude and the conduit d-c resistance. The same is true of the values obtained in tests B2 and B3 at high current. The oscillogram associated with test B8 is reproduced in Fig. 3. Notice the lag in the appearance of voltage in the circuit external to the conduit (E_{CG}). Magnetic saturation has been occurring in the steel conduit starting at the inner surface and progressing toward the outer wall. As complete saturation occurs the voltage appearing in the external circuit rises abruptly. Note that the current flow in the conduit had nearly reached crest value when this occurred.

Another interesting effect of the conduit is displayed when the conduit is not a part of the return electric circuit. Suppose that only the *H* cable is available as a return current path (all other terminals open). The current flowing out over the *A* conductor and returning over the *H* cable links the magnetic conduit as a transformer winding links its core. There tends to be reflected into the electric circuit a high magnetizing reactance modified only by the circulating current in the conduit metal. The conduit, not being laminated, allows a circulating current to flow much as a short-circuited secondary winding. This circulating current flows in a direction opposite to that in the *A* conductor on the conduit's inside surface and returns on the outside surface. In the equivalent circuit this circulating current is accounted for by a highly resistive circuit in parallel with the magnetizing reactance. The net impedance reflected into the electric circuit is much higher than would prevail if the conduit were part of the return electric circuit. These effects in iron conduits are described elsewhere.[1]

Figs. 4 and 5 give the phasor relationship of voltages and currents, together with the wave-form pattern of the voltage appearing along the exterior of the conduit as derived from several representative oscillograms.

In test B2, Fig. 4, the entire current was forced to return over the conduit alone. Note that the voltage E_{CG} appearing along the conduit exterior lags considerably behind the conduit current I_C, as was predicted. The wave shape of this voltage E_{CG} clearly shows the abrupt rise at the time of complete magnetic saturation.

In test B9, Fig. 4, the internal *B* cable was in parallel with the conduit and the return current divided nearly equally. The fact that the conduit current Ic leads the *B* cable current by a substantial angle is indicative that the conduit current is taking the highly resistive path along the inner conduit surface. The externally measured voltage E_{CG} exhibits the pattern found in test B2. After conduit saturation the external voltage closely approaches the current resistance drop computed from conduit d-c resistance.

In test B6, Fig. 5, the building structure was paralleled with the conduit. The large angle between the conduit current and the building frame current identifies the building structure as a highly inductive circuit while the conduit path appears as highly resistive. Although the current in the building frame was almost 1,100 amperes, it resulted in only a slight reduction in the conduit current because of the large phase displacement between the two current components. The voltage appearing along the conduit exterior in this test was impressed on the building structure and would appear as potential difference along beams and columns.

In test B10, Fig. 5, both the exterior grounding conductor (*H* cable) and the building frame (*G* terminal) were connected to provide parallel paths for the return current, but the conduit circuit was left open (*C* terminal not connected). The test results clearly evidence the powerful forces tending to maintain current flow in the conduit circuit. Note that across the open connection at the *C* terminal appeared a voltage of 146 volts (or, more significantly, over 50 per cent of the total impressed driving voltage), that required to force the current to return via the *H* cable and building frame in parallel. Such a voltage could be a serious shock hazard. Furthermore, unless the conduit were well insulated throughout the entire length, there would be a sufficient number of sparks at stray contact points to constitute a serious fire hazard. It was during this test that a shower of sparks observed between metallic members in the building system was caused simply by the strong magnetic field extending far out from the power conductors.

Conclusions

The significance of this investigation points unmistakably to the conclusions presented earlier in this paper. Effective use of the conduit or raceway in the equipment-grounding system is paramount. Additional work is needed to develop joints which will not throw fire during ground fault. To improve effectiveness requires greater conductivity in the conductor enclosure or the use of an internal grounding conductor. Grounding cables connecting building structure with ground electrodes (connection to earth) are needed to convey lightning currents or similar currents seeking a path to earth, but these conductors will play a negligible part in the performance of the equipment-grounding system. Of course, the importance of proper equipment-grounding becomes greater with the larger size feeder circuits and the availability of higher short-circuit currents.

Reference

1. Iron Conduit Impedance Effects in Ground Circuit Systems, A. J. Bisson, E. A. Rochau. *AIEE Transactions*, vol. 73, pt. II, July 1954, pp. 104–07.

◆

Discussion

J. A. Gienger (Eastman Kodak Company, Rochester, N. Y.): The material covered in this paper is timely. A number of cases of faulty operation in this plant have been identified with the characteristics of grounding circuits presented. We believe that there is a real need for curves or tables of volts versus amperes per unit length of conduit in the common sizes and types, and that the greatest interest would be for data on the circuit in which the current flow is *A* to *C* as illustrated in Fig. 1.

The bad showing of the conduit joints in throwing fire when conducting high current, we believe, is due entirely to improper assembly or unsuitable fittings. A number of cases of sparking at conduit couplings or terminals in this plant have come to our attention; at least one fire was started by such sparking. In every case involving couplings, however, we found the coupling loose or conclusive evidence of previous looseness.

We recently tested two 2 1/2-inch couplings at 8,500 amperes. The current was maintained for approximately two seconds. Uncleaned factory threads were employed in each case. The joints were assembled by one man utilizing 24-inch pipe wrenches. In the two couplings (four threads) tested, no sparking resulted in this tight condition. No sparking, resulted when the joints had been loosened one turn. However, in all cases, when the threads were backed off two turns, which is equivalent to "hand-tight," extensive sparking resulted. This corresponds to results of our tests made several years ago on smaller conduit couplings.

Many electricians assume that the conduit is used essentially for mechanical protection and do not realize the likelihood of sparking when the conduit is called upon to carry fault currents. We believe there are other electricians who may know better but are rather careless in their work and this results in loosely assembled conduit systems which will definitely throw sparks if they are called upon to carry fault currents.

We should be interested in a practical design of joint which an unskilled mechanic could not possibly assemble in an inadequate way. In the meantime, however, we intend to emphasize the proper assembly of our conduit system. We believe there is no practical substitute for a well-designed and well-assembled conduit system.

The inadequate current-carrying capacity of ordinary lock nuts and bushings has been recognized for many years. The National Electric Code currently specifies grounding bushings and bonding under specific situations. In our recent tests, properly bonded conduit terminals and properly assembled grounding bushings in the case of 2 1/2-inch conduit carried 8,500 amperes without enitting sparks. Again, improperly assembled terminals emitted generous showers of sparks.

Our tests were limited by our test equipment to 8,500 amperes. From the results of the tests on partly loosened couplings, a properly tightened joint would probably carry without sparking several times this value of current.

L. Brieger (Consolidated Edison Company of New York, Inc., New York, N. Y.): This excellent paper should be welcomed for its wealth of reliable information on an important subject on which very little factual material has been published.

The author demonstrates quite clearly that where an iron conduit is part of an electrical circuit, the routine notions of circuit behavior may lead to considerable error of judgment. The theoretical analysis offered in explanation of the tendency of an iron conduit to restrain the flow of return current within itself and prevent leakage to external paths is valuable indeed. Similar conditions exist in other applications of a somewhat different nature, such as, in the case of single-phase faults on pipe-type high tension cable.[1]

Although the paper is specifically addressed to the problem of equipment grounding, it may not be amiss to emphasize here that care must be exercised also in dealing with the grounding of the circuit neutral. The circuit grounding conductor is often installed in a conduit bonded to the conductor at both ends as required by the National Electrical Code. The conduit is thought of primarily as mechanical protection rather than as an element of the circuit, but a little theorizing on skin effect would indicate, and tests prove, that the current set up under accidental conditions while the grounding conductors are performing their intended protective function is carried almost entirely by the conduit. Table III shows test data on several combinations of circuit grounding conductors and conduits.

It is apparent that the conduit must be treated as a conductor, and therefore joints must be clean and tight. Mr. Kaufmann's reference to a fan-shaped shower of sparks brings out very vividly the inherent danger of poor contact. In our own tests, erratic current division between conductor and conduit was occasionally noted unless the conduit ends were cleaned before joining. During a 450-ampere test on no. 4/0 cable in $1^1/_4$-inch conduit, which was not cleaned before joining, one of the joints became blue because of high resistance at that point.

Fig. 6 shows typical impedance and voltage drop characteristics of circuit-grounding conductors in conduits. Such curves are helpful for estimating the amount of voltage rise in the grounding circuit and for other problems in connection with the application of grounds.

As proper equipment and circuit grounding is a paramount measure of safety, it is to be hoped that Mr. Kaufmann and others will keep on contributing to the understanding of this subject.

Table III. Test Data

Conductor	Conduit, Inches	Total Current, Amperes	Current in Conductor, Amperes	Current in Conduit
6	$^1/_2$	100	3	97
6	$^1/_2$	300	5	295
2	$^3/_4$	90	7	83
2	$^3/_4$	350	10	340
2/0	1	150	15	135
2/0	1	590	5	585
4/0	$1^1/_4$	225	15	210
4/0	$1^1/_4$	885	15	870

Reference

1. Single-Phase Impedance to Ground in Pipe-Type Cable, E. R. Thomas. *AIEE Transactions*, vol. 73, pt. III, Apr. 1954, pp. 336–44.

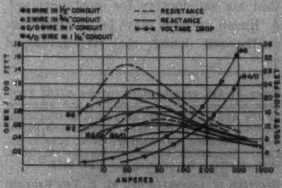


Fig. 6. Characteristics of grounding conductors. Rubber-insulated single conductor in conduit, conductor bonded to conduit at both ends. Return circuit spaced at 30 feet from grounding conductor in conduit

R. H. Kaufmann: The intense interest which has been exhibited in this paper is most gratifying and it constitutes ample reward for the work which was involved in its creation.

Mr. Gienger presents the results of tests and experience which are very interesting. His experience confirms the contention that the ground return current will, to a large extent, follow the metallic enclosure surrounding the power cables. The report on his test work indicating that adequate conductivity can be secured in threaded couplings of conduit runs is reassuring. This work indicates that screw joints, if properly cut and tightened, will probably represent no limitation in the conductivity of short-circuit currents.

Mr. Brieger brings an interesting reverse significance in the effect of a conduit surrounding a grounding conductor. In the instance reported, the outgoing current will be flowing on conductors external to the conduit rather than internal within the conduit. Hence, the grounding conductor with its enclosing conduit functions as an independent electrical conductor with the outgoing current flowing on a separate path external to the conduit. Under these circumstances, the current flow in the grounding conductor combination will encounter the lowest circuit reactance if it flows along the outside surface of the conduit. Thus, the conduit functions as a magnetic shield preventing the copper conductor within the conduit from carrying substantial current.

Mr. Brieger also observes the necessity for careful attention in joining and terminating sections of conductor enclosure to avoid thermal distress at these joints.

It is gratifying to note that the fruits of these investigations will lead directly to improved grounding techniques, better utilization of material, and to ways of avoiding application of inactive material. So far, this comprehensive subject has been treated only superficially. There remains great opportunity for fruitful investigation and it is urged that those with technical understanding and test facilities become active participants in expanding the useful knowledge relating to this field.

Measurement of Some Nonlinearities in Servomechanisms

DIETRICH K. GEHMLICH
STUDENT MEMBER AIEE

M. E. VAN VALKENBURG
ASSOCIATE MEMBER AIEE

Synopsis: A method is given for establishing the describing functions and transfer function for a servomechanism with saturation and hysteresis nonlinearities. This method makes use of a technique given by V. B. Haas, Jr. (in a personal communication to the authors) in measuring the phase and amplitude of the first harmonic of the system output with respect to a sinusoidal input.

MOST techniques for the analysis and synthesis of servomechanisms are based on the assumption that a system is linear, or nearly so. However, most elements used in servomechanisms are nonlinear to some degree in the sense that their behavior cannot be described by an ordinary linear differential equation with constant coefficients. Even so, it is fairly common practice to assume such systems are linear and to obtain results that are satisfactory for most purposes. A linear system may be analyzed or synthesized using the well-developed linear theory for the sinusoidal steady state. In contrast, the exact analysis of nonlinear systems is difficult, so that resort is often made to analogue computers to obtain usable results.

An approximate method based on the equivalent linearization of the nonlinear system has been proposed by Krylov and Bogoliuboff.[1] This method of linearization was first applied to contactor servomechanisms by Kochenburger[2] and by Goldfarb.[3] The application of equivalent linearization to nonlinear servomechanisms was further developed by Johnson.[4] Since the appearance of the Kochenburger paper in 1950, several other papers have appeared using this method.

In the work of Kochenburger and Johnson, the nonlinear system is characterized by a describing function which is analogous to a transfer function in linear theory. The describing function is found by assuming a sinusoidal input

$$x = a \sin \omega t \qquad (1)$$

and computing the output $g(x)$ which depends on the nature of the nonlinear system. The output for a sinusoidal input may be expressed as a Fourier series. The complex ratio of the first harmonic of the output to the input is defined as the describing function. The describing function is symbolized as $N(\alpha)$ where α is the normalized input magnitude

$$\alpha = \frac{a}{b} \qquad (2)$$

where

a = maximum input

b = a constant describing a particular nonlinearity

In general, $N(\alpha)$ has a magnitude and phase

$$N(\alpha) = |N(\alpha)| \exp j \operatorname{ang} N(\alpha) \qquad (3)$$

which can readily be computed. The describing function may be treated in much the same manner as a transfer function, except that the describing function depends on input amplitude as well as on frequency for some cases.

In analyzing a particular servomechanism, the linear elements are characterized by the transfer function $G(j\omega)$ and the nonlinear elements, or portions of elements, by the describing function $N(\alpha)$. The system may be studied in terms of the product $G(j\omega)N(\alpha)$ using standard steady-state methods such as the Nyquist stability criterion.

This paper is concerned with a method of measuring both the linear transfer function and the describing function of a servomechanism by means of frequency-domain measurement techniques. It will be assumed that there is access to the input and output of the system only.

The types of nonlinear elements in the servomechanism under study have been restricted to two commonly encountered in actual systems; the first is hysteresis including backlash and free play in mechanical linkages, and the

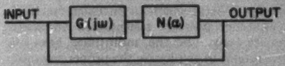


Fig. 1. Representation of nonlinear system. N(α) may represent several nonlinear units

Paper 54-285, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted January 25, 1954; made available for printing April 23, 1954.

DIETRICH K. GEHMLICH and M. E. VAN VALKENBURG are with the University of Utah, Salt Lake City, Utah.

This work was supported by the U. S. Signal Corps Engineering Laboratories, Belmar, N. J., under Contract No. *DA-36-039-SC-42621* with the Department of Electrical Engineering, University of Utah.

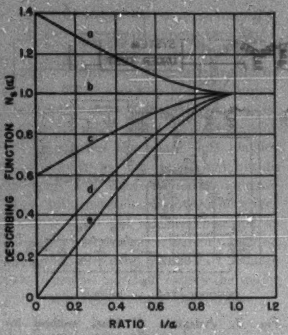


Fig. 2. Describing functions for saturation-type nonlinear elements. m = slope of characteristic curve

(A) m = 1.4
(B) m = 1
(C) m = 0.6
(D) m = 0.2
(E) m = 0.0

second is saturation or limiting. It will be assumed that there is but one dominant saturating element and one dominant hysteresis element. Further, any gain in the nonlinear systems will be assumed to be part of the linear portion of the system. A block diagram representation of the system considered is shown in Fig. 1, in which $N_s(\alpha_1)$ is the describing function for saturation and $N_b(\alpha_2)$ is the describing function for backlash (or hysteresis). The describing functions for both saturation and hysteresis have been given in the literature.[5] Figs. 2 and 3 show the results of these studies.

For the system operating open loop, the relationship between input and output is

$$\frac{\theta_o}{\theta_i} = G(j\omega)N_s(\alpha_1)N_b(\alpha_2) \tag{4}$$

and for the closed-loop system, shown in Fig. 1, the input and output are related as

$$\frac{\theta_o}{\theta_i} = \frac{G(j\omega)N_s(\alpha_1)N_b(\alpha_2)}{1+G(j\omega)N_s(\alpha_1)N_b(\alpha_2)} \tag{5}$$

The stability of the system can be studied in terms of the roots of the denominator which are defined by

$$G(j\omega) = \frac{-1}{N_s(\alpha_1)N_b(\alpha_2)} \tag{6}$$

Measurement of Nonlinearities

There are two significant differences in the describing functions for hysteresis and saturation. Because of these differences, they are separable. They can also be isolated from the linear transfer function $G(j\omega)$. A comparison of the two describing functions shows that:

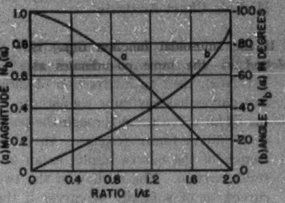


Fig. 3. Describing functions for hysteresis-type-nonlinear elements

1. The hysteresis describing function has an associated phase shift; the saturation describing function has none.
2. The effect of hysteresis is most noticeable at low input magnitudes while the effect of saturation occurs at relatively high magnitudes of input.

It should be noted also that the transfer function does not vary with input magnitude. Upon these three statements is based the method of measurement in which $G(j\omega)$, $N_s(\alpha)$, and $N_b(\alpha)$ are separated and determined. The procedure will be illustrated by a simple example.

The equipment for measurement is described in the latter part of this paper. The generator must be variable both in frequency and in magnitude, and calibrated for each. The phase angle and magnitude of the first harmonic of the output with respect to the sinusoidal input must be measured accurately. The general procedure is as follows for measurement.

1. Phase-angle measurement. At a fixed frequency, measure the phase angle of the first harmonic of the output with respect to the input as a function of the magnitude of the input for several frequencies. Such a plot is shown in Fig. 4. For any one frequency, phase changes are due to hysteresis. Hysteresis introduces a 90-degree phase lag at low magnitudes and approximately zero phase shift at large magnitudes. Thus at large magnitudes of the input, the phase shift is that of the transfer function $G(j\omega)$. Then use the phase angle of the largest input as a reference and measure some other input phase shift shown as ϕ in Fig. 4. The noise level will presumably limit the lowest magnitude that can be selected. A known ϕ for backlash fixes the magnitude of $N_b(\alpha)$ and also the $1/\alpha$ at which this ϕ occurs. Since the input a is known and the value of α has been determined, the quantity b can be computed from the equation

$$b = \frac{a}{\alpha} \tag{7}$$

This value of b is measured in terms of the system input. This constant characterizes the backlash describing function and $N_b(\alpha)$ is known.

2. Magnitude of output measurement. At a fixed frequency, record the magnitude of the first harmonic of the output as a function of the magnitude of the input. Again, this is to be repeated

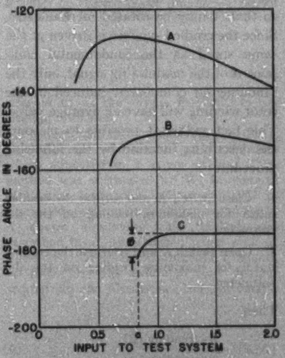


Fig. 4. System input versus phase angle

(A) Frequency = 0.5 cps
(B) Frequency = 0.75 cps
(C) Frequency = 1.0 cps

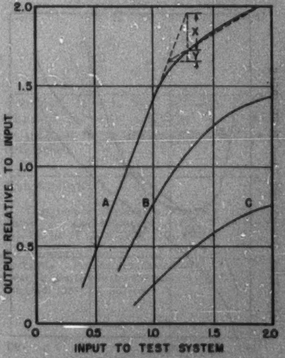


Fig. 5. System input versus output

(A) Frequency = 0.5 cps
(B) Frequency = 0.75 cps
(C) Frequency = 1.0 cps

for several different frequencies. The resulting plot is shown as Fig. 5. The effect of saturation takes place at a particular input value I_1. This input can be found graphically by extending the linear portions of the curves and sketching in the asymptotes. This determines b in the describing function $N_s(a/b)$ and so characterizes the describing function except for the value of m. To determine the quantity m, select some other input greater than I_1. At this value of input, let $X+Y$ be the incremental increase in output over that of I_1 if there was no saturation present; see Fig. 5. Also let Y be the actual increase in output at the larger input. The slope of the saturation characteristic curve is then

$$m=\frac{Y}{Y+X} \tag{8}$$

The describing function of $N_s(\alpha)$ is then defined. With the two describing functions determined, their effect can be computed at some convenient value of input, as selected by inspecting the plots made in items 1 and 2, to determine the transfer function $G(j\omega)$ for the linear portion of the system. This procedure amounts to dividing the over-all transfer function $N_T(\alpha)=G(j\omega)N_s(\alpha)N_b(\alpha)$ obtained by direct measurement, by the magnitude and phase angle of $N_s(\alpha)$-$N_b(\alpha)$. $G(j\omega)$ should remain the same for any magnitude of input. These results can then be graphed on a Nichols chart[6,7] and stability criteria and performance characteristics of the servomechanism analyzed.

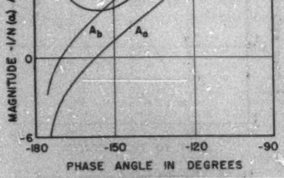


Fig. 6. Typical Nichols plot

(A) G(jω)
(B) —1/N (α)

Nichols Chart Analysis

If the sinusoidal function $G(j\omega)$ is plotted on the same co-ordinates as $-1/N(\alpha)$, then an intersection of these two curves has the same significance as the $G(j\omega)$ curve going through the point $(-1+j0)$ in linear systems. An intersection of the curves results in an unstable system. The display of the two curves, or of a family of curves if the nonlinear element is frequency sensitive, provides a means of visualizing stability for different input amplitudes. This is illustrated in Fig. 6 which shows $-1/N(\alpha)$ and $G(j\omega)$ plotted on a Nichols chart. For curve A_a, the system is stable for all values of α, although the relative stability changes with input amplitude. With the gain of the system increased to give curve A_b, the two plots intersect and instability is indicated.

Measurement of First Harmonic Amplitude and Phase

The measurement of describing functions poses a difficult problem because of the requirement that the amplitude and phase of the first harmonic be measured with respect to the sinusoidal input. To measure the first harmonic, it is necessary that all higher harmonics be rejected as completely as possible. Filters for frequencies in the range 0.1 to 10 cycles per second (cps) require prohibitively large elements. An alternate method that circumvents the filter problem has been proposed by Haas. The test setup is shown in Fig. 7. It involves two synchros driven by a variable speed motor. One of the synchros is cradled so that it may be rotated mechanically. Since the cradled synchro is driven at the same speed as the fundamental component of the modulating signal, only the fundamental component induced in the rotor winding will have an average value. This fact makes it possible to measure the describing function by the following procedure:

1. Close switch K and adjust the cradled stator for maximum reading on the d-c voltmeter.
2. Open switch K and readjust the cradled stator for maximum reading on the d-c voltmeter.

Then

$$|N(\alpha)|=\frac{\text{d-c reading (2)}}{\text{d-c reading (1)}} \tag{9}$$

and

$$\text{ang } N(\alpha)=\text{difference in cradle position (1) and (2)} \tag{10}$$

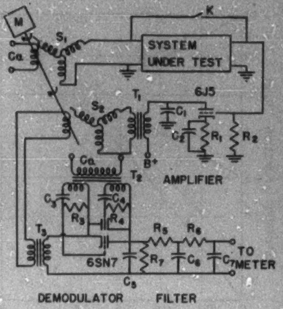


Fig. 7. Adaption of Haas method for measuring describing functions

M = variable speed motor
C_a = carrier signal
S_1 = 400-cycle synchro
S_2 = cradled synchro
T_1, T_2 = matching transformers
T_3 = primary—115 volts, secondary—6.3 volts
C_1, C_3, C_4 = 0.05 microfarad
C_2 = 2 microfarads
C_5 = 1 microfarad
C_6, C_7 = 16 microfarads
R_1 = 2,700 ohms
R_2, R_7 = 1 megohm
R_3, R_4 = 500,000 ohms
R_5, R_6 = 27,000 ohms

The purpose of the amplifier is to isolate the cradled synchro input such that no loading of the system under test will occur. This unit can thus be made to test any system provided that the signal from the exciting synchro is sufficient to drive the servomechanism under test.

Discussion

The question will arise as to how good an approximation will be given by the analysis of a nonlinear system with describing functions. The answer is difficult because it depends on the particular system under analysis. The method ignores harmonics other than the first. Many systems contain elements that behave as low-pass filters, e.g., motors, motor fields, mechanical components. These elements filter higher harmonics so that they are attenuated before returning around the loop to the input. This frequency-dependent attenuation is the key to the success of the method. In general, it can be said that the more complex the system, in the sense that it contains many low-pass filter components appearing as lag terms in the transfer function, the better will be the approximation.

There are indications that the measurement method can be applied to other types of nonlinearities. Haas[8] has shown that the describing function for coulomb friction is dependent on both the amplitude and frequency of the input signals. A study of Figs. 4 and 5 shows that at low frequencies the phase shift reaches a minimum as the input magnitude increases and then begins to increase again with increasing input. This effect becomes more and more pronounced as frequency decreases. Coulomb friction causes the increasing phase shift at large inputs just as backlash causes the major phase shift at low input magnitudes.

Application to a Typical Servomechanism

A positional servomechanism designed to be as nearly linear as possible was built and tested. To this servomechanism were added experimental backlash and saturation units for which the describing constants were known. The assumption was made that the added nonlinear units would be of such magnitude that any inherent nonlinearity in the servo could be neglected. A test of this system was made to check the general procedure outlined and the accuracy of the method. In this example the drift was negligible and the describing function data were obtained with the system operating open loop.

The nonlinear units added to the system were adjusted to have the following characteristics:

1. Hysteresis: $b_0=0.29$.
2. Saturation: $b_s=1.0$, $m=0.23$.

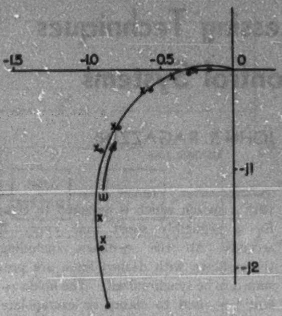


Fig. 8. Complex plane plot of G(jω) for test system

(x)—G(jω) for a = 1.5
(o)—G(jω) for a = 1.0

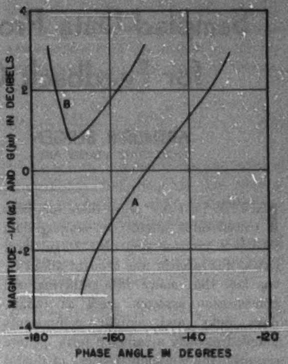


Fig. 9. Nichols plot for test system

(A) G(jω)
(B) −1/N(α)

Following the procedure given, characteristics of the system were obtained as follows:

1. For $N_s(\alpha)$: $b_s=1.1$ from Fig. 5, and $m=1.4/6.3=0.22$.
2. For $N_b(\alpha)$: at a frequency of 1 cps and with $a=0.83$, θ due to backlash $=-10$ degrees from Fig. 4. From Fig. 3, for $\theta=-10$ degrees, $1/\alpha=0.33=b/a=b/0.83$, $b=0.275$.

Knowing $N_s(\alpha)$ and $N_b(\alpha)$, $G(j\omega)$ may be found. The complex plane plots of $G(j\omega)$ for this example are illustrated in Figs. 8 and 9 for two input amplitudes.

References

1. Nonlinear Mechanics (book), **N. Minorsky.** Edward Brothers, Ann Arbor, Mich., section by Krylov and Bogoliuboff, 1947.

2. A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms, **R. J. Kochenburger.** *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

3. On Some Nonlinear Phenomena in Regulatory Systems (in Russian), **L. C. Goldfarb.** *Avtomatika i Telemekhanika*, Moscow, Russia, vol. 8, no. 5, 1947, pp. 349–83. Translated in *Report 1691*, National Bureau of Standards, Washington, D. C., 1952.

4. Sinusoidal Analysis of Feedback-Control Systems Containing Nonlinear Elements, **E. Calvin Johnson.** *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 169–81.

5. Describing Function Method of Servomechanism Analysis Applied to Most Commonly Encountered Nonlinearities, **H. D. Greif.** *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 243–48.

6. Servomechanisms and Regulating System Design, **Harold Chestnut, Robert W. Mayer.** John Wiley and Sons, Inc., New York, N. Y., 1951.

7. Theory of Servomechanisms, **Hubert M. James, Nathaniel B. Nichols, Ralph S. Phillips.** McGraw-Hill Book Company, Inc., New York, N. Y., 1947.

8. Coulomb Friction in Feedback Control Systems, **Vinton B. Haas, Jr.** *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 119–26.

◆

Discussion

Vinton B. Haas, Jr. (Massachusetts Institute of Technology, Cambridge, Mass.): The authors are to be complimented for their paper on the measurement of some nonlinearities. I am pleased that the information contained in my letter to Prof. Van Valkenburg was useful to them in their research; however, I would like to point out that the idea of the technique for measuring the fundamental component referred to as the "Haas method" is not original with me.

I learned of this method from Dr. R. J. Kochenburger who used it to obtain data which appeared in a paper of his.[1] The method was originated by W. Learmonth of Vickers-Armstrong Ltd. in mid-1947 in connection with some nonlinear servo work. A device employing this method was described in an unpublished paper read at a servo conference in Birmingham, England, in 1950. It was exhibited and demonstrated at the Conference on Automatic Control.[2] It was also exhibited at the Physical Society's exhibition in London in 1952, and was briefly described in the catalogue.

REFERENCES

1. Limiting in Feedback Control Systems, **Ralph J. Kochenburger.** *AIEE Transactions*, vol. 72, pt. II, July 1953, pp. 180–94.

2. Automatic and Manual Control, edited by **A. Tustin.** *Proceedings*, Conference on Automatic Control, Butterworths Publications Ltd., London, England, 1951, pp. 567–68.

Sampled-Data Processing Techniques for Feedback Control Systems

ARTHUR R. BERGEN
ASSOCIATE MEMBER AIEE

JOHN R. RAGAZZINI
MEMBER AIEE

IN THE past few years there has been considerable interest in sampled-data feedback control systems.[1-6] In most cases, this interest was brought about by the fact that many data-gathering and transmission systems, such as radars, data links and similar components handled data in a sampled form. More recently, the contemplated use of digital computers, in essence sampled-data devices, in connection with process control, has necessitated some study of means for utilizing their capacity to compute for purposes of stabilizing and shaping control systems.[7,8] By their nature, digital computers are sampled-data systems since their output is intermittent and in numerical form.

Several problems arise in sampled-data systems which are unique to such systems and which tend to limit their performance and applicability. These include a tendency toward instability, ripple components at sampling frequency and its harmonics in the output, and a restriction in usable bandwidth to a fraction of sampling frequency. Yet for certain applications there are advantages to stabilization techniques employing the data samples directly which outweigh the disadvantages to an extent that makes sampled-data control systems desirable even when not necessary by the nature of the input to the system. The applications in which sampled-data processing units can most profitably be employed are those involving large-scale process controls, industrial drives, regulators, etc., where the additional complexity necessary to mechanize these units is not a burden. In some cases, implementation can take the form of digital computers, while in others analogue devices are more desirable. In the material which follows, the theory, design, limitations, and applications of sampled-data processing systems are given.

Description of System

In the first place, the sampled-data feedback control system, considered in the following discussions, is assumed to be error-sampled as shown in Fig. 1. The process of sampling is shown symbolically with a switch which is assumed to close for a negligibly short time every T seconds. All the switches, including those shown with dashed lines, are presumed to be synchronized. The hold system[1,3] is used to clamp or extrapolate the data pulse for the interval between samples. This element serves to improve the duty cycle of the plant as well as to reduce the amount of ripple in the controlled variable C. If there were no processing unit present, the sampled error would be applied directly to the forward transmission system G. Stabilization and shaping, by the use of continuous networks, suffer limitations, in part because of the inherent limitations of the control functions which can be obtained.

The processing unit, shown in Fig. 1, receives the error samples, E_1^*, stores a certain number of them, does the same with a number of past samples of its output, E_2^*, weights them in a specified way and delivers to the system a modified error pulse E_2^*. By properly choosing the weights given to each of these samples, the modified pulse train, as computed by the processing unit, yields some desired over-all performance by the system. As will be developed later, certain optimum types of performance are obtainable with systems like these. Further, the processing unit will not suffer from the complexity of being required to store an inordinate number of data samples. In addition, the time scale of most process control systems is such that relatively simple analogue storage devices can be employed to advantage.

General Background

It has been shown that in dealing with sampled-data systems, the z-transform method can be readily used for design.[3-6] This method is essentially a Laplace transform technique in which sampled pulse trains are related by system functions in the form of z-transforms. These system functions are ratios of rational polynomials in the variable z where z is defined as ϵ^{Ts}. It has been shown[3] that if two systems are cascaded and separated by a synchronous switch, the over-all system function is simply the product of the two system functions. This relationship is fortuitously convenient but is not the only reason why sampled-data processing techniques are useful or desirable. The main reason for considering them is that by storing past samples of its input and output pulse sequences rational polynomial ratios needed for stabilization and shaping can be obtained readily.

Stated more simply in the time domain the processing unit solves the relationship which appears in the form of a difference equation relating the past input and output samples as follows

$$\begin{aligned} a_0e_1[(nT)]+a_1e_1[(n-1)T]+a_2e_1[(n-2)T]+ \\ \ldots a_je_1[(n-j)T]=b_0e_2[(nT)]+ \\ b_1e_2[(n-1)T]+\ldots b_ke_2[(n-k)T] \end{aligned} \tag{1}$$

where $e_1(nT)$ represents the nth sample of the input error $e_1(t)$ as shown in Fig. 2. Applying the z-transformation to both sides of equation 1, there results

$$\begin{aligned} [a_0+a_1z^{-1}+a_2z^{-2}+\ldots a_jz^{-j}]E_1^*(z)= \\ [b_0+b_1z^{-1}+b_2z^{-2}+\ldots b_kz^{-k}]E_2^*(z) \end{aligned} \tag{2}$$

which yields for the system function describing the processing unit

$$D^*(z)=\frac{E_2^*(z)}{E_1^*(z)}=\frac{a_0+a_1z^{-1}+a_2z^{-2}+\ldots a_jz^{-j}}{b_0+b_1z^{-1}+b_2z^{-2}+\ldots a_kz^{-k}} \tag{3}$$

It is observed that by taking arbitrary weights for $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$, any polynomial ratio in z^{-1} can be obtained. These polynomials are not subject to the restrictions pertaining to those obtainable with a pulsed network in place of the processing unit. The only restriction on physical realizability is that no future samples, but only present and past samples, can be considered in realizing the system function of the processing unit, $D^*(z)$. This implies that when $D^*(z)$ is expressed as a ratio of polynomials in z, the order of z in the numerator should be no higher than that of the denominator. In addition, it can readily be seen that the constant b_0 must always be present to meet this condition.

It has been shown[3] that the input-output relation for the error-sampled system shown in Fig. 1 is given by

$$K^*(z)=\frac{C^*(z)}{R^*(z)}=\frac{D^*(z)G^*(z)}{1+D^*(z)G^*(z)} \tag{4}$$

Paper 54-282, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted December 30, 1953; made available for printing May 6, 1954.

ARTHUR R. BERGEN and JOHN R. RAGAZZINI are with Columbia University, Electronics Research Laboratories, New York, N. Y.

This research was supported by the United States Air Force under Contract No. *AT 18(600) 677* monitored by the Office of Scientific Research, Air Research and Development Command.

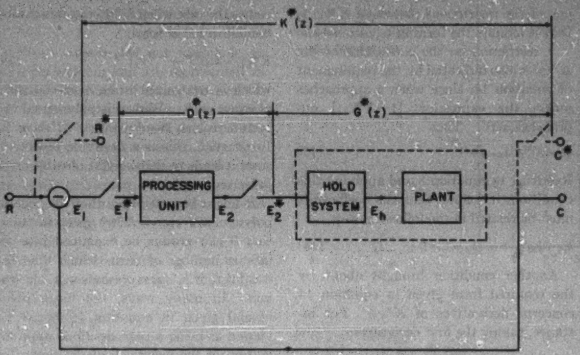


Fig. 1. Typical sampled-data control system using data processing unit

The fact that the z-transform of the forward transfer function, $D^*(z)G(z)$, is simply the algebraic product of the stabilizer system function $D^*(z)$ and the plant system function $G^*(z)$ is a fortuitous fact which makes the system-shaping a problem no more difficult than that for continuous linear systems. Thus, the design of the sampled-data processing unit becomes that of choosing the function $D^*(z)$ in such a manner as to give a prescribed over-all performance to specified inputs.

This choice is not only limited by the physical realizability conditions listed in the previous paragraph but also is subject to certain restrictions which will be developed later in this paper. Previous work in this field[6,8] has been directed either to the shaping of the internal transfer function of the sampled-data system in the frequency domain using processing units implemented in the form of digital computers, or to the adjustment of the roots of the characteristic equation of the over-all transmission, by means of a stabilizer which is similar to the sampled-data processing unit. By contrast, the techniques developed here are used to synthesize systems which will respond in a prescribed manner to a test input or set of inputs in the time domain. The frequency response characteristics become incidental, though they are readily obtainable if desired.

Prototype Forms for K*(z)

1. Limitations of Physical Realizability

The limitations which are imposed on the forms of $G^*(z)$ and $D^*(z)$ by their physical realizability restrict the choices of possible over-all transmissions $K^*(z)$. Taking first the restrictions placed on $G^*(z)$, it is assumed that the plant and associated hold system is one whose step response is roughly of the form given in Fig. 3. The plant transfer lag ζ is zero in many cases. The response may be monotonic or oscillatory but in all cases the step response starts from zero. This implies that no poles of $G(s)$ lie in the right half of the complex plane and that the degree of the denominator of $G(s)$ is higher than that of the numerator. Generally speaking, the pulsed transfer function $G^*(z)$ is the ratio of two rational polynomials with the degree of the denominator higher than that of the numerator. Typically

$$G^*(z)=\frac{p_m z^{-m}+\ldots p_a z^{-a}}{q_0+q_1 z^{-1}+\ldots q_b z^{-b}} \tag{5}$$

where m is the number of the first interval immediately following the lag time ζ and q_0 is a constant which is always present in physically realizable systems.

Because of these aforementioned properties, $G^*(z)$ can always be expanded into an infinite series in z^{-1} as follows

$$G^*(z)=g_m z^{-m}+g_{m+1} z^{-(m+1)} \tag{6}$$

If the delay ζ is zero, the first term in equation 6 is $g_1 z^{-1}$, which indicates that the system response has a value other than zero only after one sample period has passed.

Keeping in mind these restrictions on $G^*(z)$ and those mentioned previously on $D^*(z)$, prototype over-all response functions $K^*(z)$ which can be obtained will be developed. Substituting the polynomials given in equations 3 and 6 in the expression for over-all response given in equation 4, there results

$$K^*(z)=\frac{[a_0+a_1 z^{-1}+\ldots a_j z^{-j}][p_m z^{-m}+\ldots p_a z^{-a}]}{[b_0+b_1 z^{-1}+\ldots b_k z^{-k}][q_0+q_1 z^{-1}+\ldots q_b z^{-b}]+[a_0+a_1 z^{-1}+\ldots a_j z^{-j}]\times[p_m z^{-m}+\ldots p_a z^{-a}]} \tag{7}$$

which can be simplified to

$$K^*(z)=\frac{k_m z^{-m}+\ldots k_p z^{-p}}{l_0+l_1 z^{-1}+\ldots l_q z^{-q}} \tag{8}$$

where it is seen that l_0 must be present since b_0 and q_0 are always present. For cases where the plant transfer lag ζ is zero, the over-all response becomes

$$K^*(z)=\frac{k_1 z^{-1}+\ldots k_p z^{-p}}{l_0+l_1 z^{-1}+\ldots l_q z^{-q}} \tag{9}$$

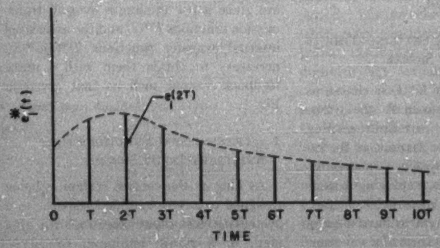


Fig. 2. Error pulse train

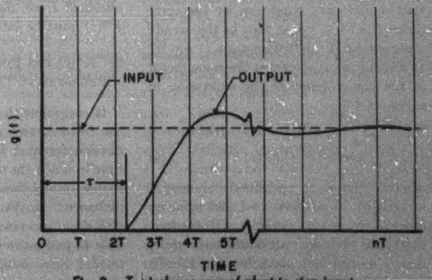


Fig. 3. Typical response of plant to step input

Thus, the classes of over-all response functions which can be obtained with the restrictions of physical realizability placed on $D^*(z)$ and $G^*(z)$ are those given by equations 8 and 9. If no plant transfer lags are present, over-all transmission functions given by equation 9 are realizable. However, there are still additional restrictions on the coefficients of $K^*(z)$ brought about by the fact that certain prescribed performance is required of the system in responding to specified input test functions. The usual functions used for this purpose are unit steps or unit ramps or, more generally, functions whose Laplace transform is $1/s^m$. The conditions imposed on the coefficients of $K^*(z)$ by the requirement that the control system follow such inputs with zero error will now be reviewed.

Referring to Fig. 1, the error sequence $E_1^*(z)$ is given by

$$E_1^*(z)=R^*(z)-C^*(z) \qquad (10)$$

which is

$$E_1^*(z)=R^*(z)[1-K^*(z)] \qquad (11)$$

It is desired, among other things, that the final value of this error sequence be zero. Applying the final value theorem[5] to equation 11 and equating the result to zero

$$E_1(nT)|_{n\to\infty}=\{R^*(z)[1-K^*(z)]\times [1-z^{-1}]\}|_{z=1}=0 \qquad (12)$$

From a consideration of the expressions relating the Laplace transform of continuous functions and the corresponding pulsed transfer functions,[3] it can readily be shown that the number of poles of $R^*(z)$ present at $z=1$ is equal to the number of poles at the origin of $R(s)$. Thus, for the test functions under consideration, equation 12 can be written

$$E_1(nT)|_{n\to\infty}=\left\{\frac{F_1^*(z)}{(1-z^{-1})^m}\times [1-K^*(z)][1-z^{-1}]\right\}_{z=1}=0 \qquad (13)$$

where $F_1^*(z)$ is an undefined polynomial in z^{-1} with no zeros at $z=1$. It is seen that a necessary condition for the error pulse $E(nT)$ to go to zero in the steady state is that $[1-K^*(z)]$ contain $(1-z^{-1})^m$ as a factor so as to cancel the same factor in the denominator of the expression in equation 13.

Hence

$$[1-K^*(z)]=[1-z^{-1}]^m F_2^*(z) \qquad (14)$$

where $F_2^*(z)$ is an undefined polynomial in z^{-1} with no zeros at $z=1$ and m is the order of the pole of s in $R(s)$. If, further, $F_2^*(z)$ has no poles outside the unit circle, this is a sufficient condition for $E_1(nT)$ to go to zero in the steady state. Thus one of the restrictions placed on $K^*(z)$ is that it assume the form in equation 14.

A restriction on the coefficients of the $K^*(z)$ is also indicated by the requirement of equation 14, since when z approaches unity, the expression $[1-K^*(z)]$ approaches zero. Thus

$$[1-K^*(z)]_{z=1}=0 \qquad (15)$$

Referring to equation 8 and applying this condition, it follows that the coefficients must be related in the following way

$$k_m+k_{m+1}+\ldots k_p=l_0+l_1+\ldots l_q \qquad (16)$$

Another condition brought about by the required form given in equation 14 concerns derivatives of $K^*(z)$. For instance, taking the first derivative

$$\left.\frac{dK^*(z)}{dz^{-1}}\right|_{z=1}=[(1-z^{-1})^m F_2'^*(z)- m(1-z^{-1})^{m-1}F_2^*(z)]_{z=1}=0 \qquad (17)$$

it is observed that as z is made equal to unity this derivative goes to zero. This condition results in still another relation between coefficients obtained by differentiating the expression in equation 8 and combining with the relation given in equation 16. This results in

$$mk_m+(m+1)k_{m+1}+\ldots pk_p=l_1+2l_2+\ldots ql_q \qquad (18)$$

It is evident by further differentiating $K^*(z)$ that as high an order of derivative as the order of the pole of the test function $R(s)$ at the origin must equal zero as z is made equal to unity. Thus, the requirement that the system follow a test function of unit step, ramp, or acceleration type without steady-state error places restrictions on $K^*(z)$ in which the coefficients must be related as given in equations 16 and 18 or similar relations brought about by equating higher order derivatives of $K^*(z)$ to zero for z equal to unity. For instance, for a unit step, only equation 16 need be satisfied and for a unit ramp both equations 16 and 18 must be satisfied. For a unit acceleration, equations 16 and 18 as well as that obtained from the second derivative of $K^*(z)$ must be satisfied.

2. $K^*(z)$ for Finite Settling Time in Response to Test Signal

It is evident that if $K^*(z)$ is chosen as the ratio of two polynomials the transmission function has an infinite settling time in much the same manner as the response of any linear system to a step disturbance. In systems employing a sampled-data processor, however, it is possible to cause the system to have a finite settling time so far as the values at sampling instants are concerned. This becomes the case when $K^*(z)$ is a finite polynomial in z^{-1} of form

$$K^*(z)=k_m z^{-m}+\ldots k_p z^{-p} \qquad (19)$$

which is recognized to be a special form of equation 8. With this polynomial the system reaches steady-state in pT seconds. In practice, this is a desirable objective, since it leads to the shortest possible error pulse sequences.

It might be noted here that the general polynomial form of $K^*(z)$ given in equation 9 can always be expanded into an infinite number of terms which tend toward zero if $K^*(z)$ represents a stable system. In many ways, the finite polynomial given in equation 19 might be viewed as being a very desirable approximation of the general form consisting of an infinite number of terms. The general conditions restricting the coefficients of the terms in $K^*(z)$ hold, of course, for this form also, leading to the expressions

$$k_m+k_{m+1}+\ldots k_p=1 \qquad (20)$$

$$mk_m+(m+1)k_{m+1}+\ldots pk_p=0 \qquad (21)$$

3. Prototype Transmission Functions with Finite Settling Time

In addition to the restrictions placed on the coefficients of the terms in $K^*(z)$, there are other less obvious ones imposed by power limits of the plant, saturation effects, and system stability. For certain linear systems, however, certain idealized or minimum settling time functions for $K^*(z)$ become feasible and will be studied here.

First it is noted that to follow without error an input function whose Laplace transform is $1/s^m$, there are m conditions of the type given by equations 20 and 21. Hence, no transmission function $K^*(z)$ can be achieved with less than m terms since there are m conditions which must be simultaneously satisfied. A polynomial having m terms is then referred to as a "minimum" polynomial and represents the minimum number of sample periods in which a system can be made to settle on a given test input. In Table I are given a few minimum over-all transmission functions $K^*(z)$ and the associated internal transfer functions $D^*(z)G^*(z)$, necessary to obtain them with a unity feedback system such as that given in Fig. 1.

4. Transmission Functions for Arbitrary Input Signals

As long as the control system is being designed to have a specified response to a systematic test input function such as a step, or a ramp, minimum prototype transmission functions having finite settling time can be specified as shown in the previous section. On the other hand, if the input signals are not amenable to such specification and are represented as arbitrary functions of time, the over-all performance of the system is not as readily specified. This section will deal with the performance of systems which have been designed to handle certain systematic inputs when subjected to arbitrary input functions. For simplicity, only systems having no plant process lags will be considered here.

In this case it can readily be verified that the prototype transmission $K^*(z)$ is given by

$$K^*(z)=z^{-1}\sum_{i=0}^{(p-1)}[1-z^{-1}]^i \qquad (22)$$

Thus, the output pulse sequence $C^*(z)$ resulting from an applied input pulse sequence $R^*(z)$ will be

$$C^*(z)=z^{-1}\sum_{i=0}^{(p-1)}[1-z^{-1}]^i R^*(z) \qquad (23)$$

Replacing the summation by its equivalent expression in closed form

$$C^*(z)=\frac{1-(1-z^{-1})^p}{1-(1-z^{-1})}z^{-1}R^*(z) \qquad (24)$$

Simplifying

$$C^*(z)=[1-(1-z^{-1})^p]R^*(z) \qquad (25)$$

From which the error expressed as a z-transform becomes

$$E^*(z)=R^*(z)-C^*(z)=(1-z^{-1})^pR^*(z) \qquad (26)$$

Taking the real inversion of this transform and writing the error pulse sequence

$$E(nT)=\nabla^pR(nT) \qquad (27)$$

where $\nabla^pR(nT)$ is the pth back difference[9] of $R(nT)$ and p is the number of terms in the prototype function of equation 22.

Table I. Transfer Functions Corresponding to Prototype Over-all Transmissions

Over-all Transmission $K^*(z)$	Transfer Function $D^*(z)G^*(z)$
For Step Input Test Function, $R(s)=1/s$	
z^{-1}	$\frac{z^{-1}}{1-z^{-1}}$
z^{-2}	$\frac{z^{-2}}{(1-z^{-1})(1+z^{-1})}$
z^{-3}	$\frac{z^{-3}}{(1-z^{-1})(1+z^{-1}+z^{-2})}$
z^{-4} etc.	$\frac{z^{-4}}{(1-z^{-1})(1+z^{-1}+z^{-2}+z^{-3})}$
For Ramp Input Function, $R(s)=1/s^2$	
$2z^{-1}-z^{-2}$	$\frac{2z^{-1}-z^{-2}}{(1-z^{-1})^2}$
$3z^{-2}-2z^{-3}$	$\frac{3z^{-2}-2z^{-3}}{(1-z^{-1})^2(1+2z^{-1})}$
$4z^{-3}-3z^{-4}$ etc.	$\frac{4z^{-3}-3z^{-4}}{(1-z^{-1})^2(1+2z^{-1}+3z^{-2})}$
For Acceleration Input Function, $R(s)=1/s^3$	
$3z^{-1}-3z^{-2}+z^{-3}$ etc.	$\frac{3z^{-1}-3z^{-2}+z^{-3}}{(1-z^{-1})^3}$

It is readily seen that for inputs such as ramps or steps, applied to the appropriate prototype, higher order differences become zero in the steady state thus resulting in zero error. During the transient period after the sudden application of such inputs, however, the higher order differences are not zero until a certain number of sampling intervals have passed. For example, in the case of a ramp input, two sampling intervals must pass before the second order backward difference becomes zero.

The representation of the error sequence in the form given in equation 27 can be used to extend the analysis to arbitrary input functions also. The result in equation 27 is general and can be applied to any input function. Its only restriction is that it applies to no lag systems whose transmission functions are of the type listed in Table I. An interesting and pertinent phenomenon is observed in relation to the higher order differences for arbitrary input functions. Generally speaking a "slowly varying" function is one whose higher order differences will be progressively smaller in magnitude. Such a characteristic is found in functions which do not vary much in one sampling period. On the other hand, for functions which vary considerably in one sampling period, the higher order differences do not diminish but actually increase in most cases. In such a case, the higher the order of the transmission function $K^*(z)$, the larger will the error sequence be in view of equation 27.

An example of this phenomenon can be pointed out more specifically with respect to a sinusoidal function. It can be shown by numerical methods that, so long as the frequency of the sinusoidal input is less than 1/6 the sampling frequency, the higher order, backward differences keep diminishing progressively in magnitude. Such an input will be passed with diminishing error as the order of the transmission function selected is increased. It should be pointed out, however, that this higher order function will also produce a longer settling time under transient conditions. On the other hand, if the frequency of the sinusoidal input function is higher than 1/6 the sampling frequency, the higher order differences increase progressively in magnitude. In this case, higher order transmission functions are undesirable and should be avoided.

Generalizing this characteristic relating to sinusoidal inputs, it might be stated that if the arbitrary input function can be represented during any given interval by a sinusoid or group of sinusoids whose frequency is less than 1/6 the sampling frequency, reduction of the magnitude of error sequences will be achieved by use of higher order transmission functions. On the other hand, such transmission functions also result in correspondingly longer settling times when subjected to sudden inputs. A design compromise must be found for each specific case.

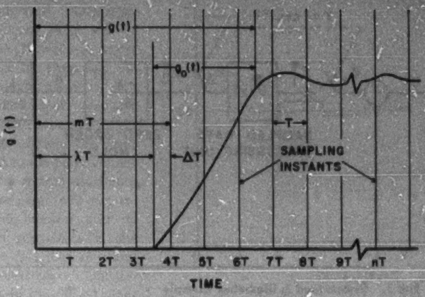


Fig. 4. Impulsive response of plant with transfer lag

Systems with Plant Transfer Lag

As stated previously, one of the major applications of sampled-data processing units is in systems having slow response such as large-scale process control. Such systems are often characterized by a plant transfer lag which is not necessarily equal to an integral number of sampling intervals. Thus, a system of this type will be described by a transfer function $G(s)$ which includes a multiplicative term $\epsilon^{-\lambda Ts}$ where λ is not necessarily an integer. To be able to design sampled-data processing units for such systems it is necessary to obtain the pulsed transfer function, $G^*(z)$ which takes into account a plant transfer lag.[6]

A typical impulsive response function

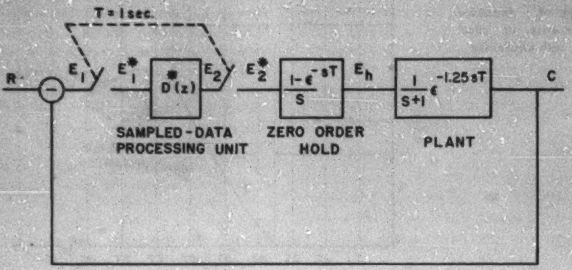

T = 1 sec.
R
E_1
E_1^*
$D^*(z)$
E_2
E_2^*
$\frac{1-\epsilon^{-sT}}{s}$
E_h
$\frac{1}{s+1}\epsilon^{-1.25sT}$
C
SAMPLED-DATA PROCESSING UNIT
ZERO ORDER HOLD
PLANT

Fig. 5. System used in illustrative example

for a plant of this type is shown in Fig. 4. The plant transfer lag is λT and

$$m=\lambda+\Delta \tag{28}$$

The constant m is an integer representing the delay in the pulsed transfer function, and Δ is a fraction whose significance is shown in Fig. 4. The usual system function for a plant of this type can be represented by

$$G(s,\lambda)=G_0(s)\epsilon^{-\lambda Ts}=\epsilon^{-mTs}G_0(s)\epsilon^{\Delta Ts} \tag{29}$$

where $G_0(s)$ is a ratio of rational polynomials in s. The z-transform for a system function of this type becomes

$$G^*(z,\lambda)=z^{-m}G^*(z,-\Delta) \tag{30}$$

In view of the fact that m is an integer, it is possible to extract it as z^{-m}. However, in evaluating $G^*(z,-\Delta)$ care must be exercised since the delay $-\Delta$ is always a fraction of s sampling interval.

Returning to the basic definition of the z-transform[3]

$$G^*(z,-\Delta)=\sum_{n=0}^{\infty} g_0(n+\Delta)Tz^{-n} \tag{31}$$

it is observed that the samples of $g_0(t)$ are taken not at sampling instants as measured from the origin of $g_0(t)$ but instants which are displaced by a fractional interval as seen in Fig. 4. The transform can be expressed in closed form by applying contour integration[5] yielding the following general expression

$$G^*(z,-\Delta)=\sum_{\text{poles of } G_0(s)}\times \text{Residues}\left\{G_0(s)\frac{\epsilon^{\Delta sT}}{1-\epsilon^{+sT}z^{-1}}\right\} \tag{32}$$

A brief listing of such transforms appears as Table II, where it is observed that as the advance Δ is reduced to zero the z-transforms reduce to ordinary z-transform pairs. Despite its brevity, this table can be used to evaluate more complex forms, since the Laplace transforms representing the continuous system function can be broken up into partial fractions and evaluated part by part.

Table II. z-Transform Pairs for Systems with Transfer Lags

$F(s)\epsilon^{\Delta Ts}$	$F(z,\Delta-)$
$\frac{1}{s}\epsilon^{\Delta Ts}$	$\frac{1}{1-z^{-1}}$
$\frac{1}{s^2}\epsilon^{\Delta Ts}$	$\frac{\Delta T+Tz^{-1}(1-\Delta)}{(1-z^{-1})^2}$
$\frac{1}{s+a}\epsilon^{\Delta Ts}$	$\frac{\epsilon^{-a\Delta T}}{1-\epsilon^{-aT}z^{-1}}$
$\frac{a}{s(s+a)}\epsilon^{\Delta Ts}$	$\frac{1-\epsilon^{-a\Delta T}+z^{-1}(\epsilon^{-a\Delta T}-\epsilon^{-aT})}{(1-z^{-1})(1-\epsilon^{-aT}z^{-1})}$

ILLUSTRATIVE EXAMPLE

To illustrate the various points brought out in the preceding theoretical discussions, a simple illustrative example will be given. The object is to demonstrate the design of a sampled-data processing unit which will cause a system having a plant transfer lag to conform to some desired over-all transmission function $K^*(z)$. The system under consideration is shown in Fig. 5, and the response of the plant to a unit step function is shown in Fig. 6. It is observed that the various constants previously described are as follows: $\lambda=1.25$ seconds, $\Delta=0.75$ second and $m=2$. The z-transform of the combined zero-order hold circuit and plant are found by use of Table II. The forward transfer function thus becomes

$$G^*(z)=\frac{(0.528+0.104z^{-1})z^{-2}}{(1-0.368z^{-1})} \tag{33}$$

If a unit step input is applied, the z-transform of the output $C^*(z)$ is given by

$$C^*(z)=\frac{(0.528+0.104z^{-1})z^{-2}}{(1-z^{-1})(1-0.368z^{-1}+0.528z^{-2}+0.104z^{-3})} \tag{34}$$

which, upon expansion by division of the denominator into the numerator gives an output sequence

$$C^*(z)=0.528z^{-2}+0.826z^{-3}+0.655z^{-4}+0.378z^{-5}+0.335z^{-6}+\ldots \tag{35}$$

The output converges to a final value of 0.500. The sampled response of the system is lightly damped and has a bad overshoot as will be seen in Fig. 7.

To demonstrate the application of a sampled-data processor it will be specified that the desired over-all response should be a simple delay. For an input step function, this would cause the response to appear as shown in Fig. 7. From Table I, the required transfer function becomes

$$D^*(z)G^*(z)=\frac{z^{-2}}{1-z^{-2}} \tag{36}$$

Therefore, $D^*(z)$, the component contributed by the stabilizer is

$$D^*(z)=\frac{D^*(z)G^*(z)}{G^*(z)}=\frac{1-0.368z^{-1}}{(0.528+0.104z^{-1})(1-z^{-2})} \tag{37}$$

As will be shown later, a physical device which produces this pulsed response can be implemented. Referring to equations 1 through 3, the "program" followed by the stabilizer is simply

$$E_2(nT)=1.89E_1(nT)-0.697E_1(n-1)T-0.198E_2(n-1)T+E_2(n-2)T-0.198E_2(n-3)T \tag{38}$$

It will be necessary to store three past

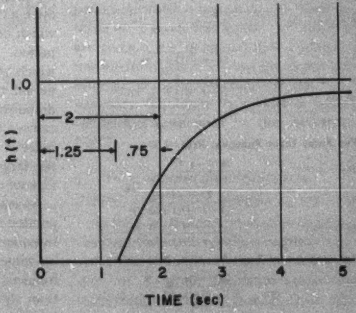


Fig. 6. Step response of plant used in illustrative example

samples of the output sequence and one past sample of the input sequence. It is interesting to observe the form of the error sequence E_2^* resulting from the application of a unit step. The z-transform of this sequence is

$$E_2(nT)=\frac{D^*(z)R^*(z)}{1+D^*(z)G^*(z)}$$
$$=\frac{1-0.368z^{-1}}{(0.528+0.104z^{-1})(1-z^{-1})} \quad (39)$$

which, upon expansion, yields

$$E_2^*(z)=1.89+0.821z^{-1}+1.03z^{-2}+0.990z^{-3}+\ldots \quad (40)$$

When the sequence represented by equation 40 passes through the zero-order hold (clamp) circuit, it yields a time function shown in Fig. 8. This function is the correct input to the plant to result in performance having no overshoots. There are some overshoots in the continuous output between sampling intervals but these are small and can be computed using one of a number of techniques.

Practical Limitations in Applying Sampled-Data Processing Units

In theory, it is always possible to design a sampled-data processing unit to synthesize an internal transfer function which will result in an over-all transmission by the prototype functions with a finite settling time. There are a number of practical considerations, however, which may make the realization of such functions impossible or impractical. In these cases, the minimum transmission functions can be used as design objectives and may be approached in varying degree with certain systems, depending on the importance of the physical limitations. The two major limitations which will be considered here are the physical one represented by saturation in the plant and the mathematical one brought about by certain improper cancellations of poles and zeros by the sampled-data processing unit. To restate the first restriction, the input error sequence applied to the plant cannot be allowed to exceed a certain maximum value. If it does, the plant will not respond linearly and will, in effect, limit its output. The second restriction is one brought about by the fact that the sampled-data processing unit cancels out undesired poles or zeros in the pulsed plant transfer function and replaces them with desired poles and zeros. The question arises as to what will be the effect of a slight shift in the poles and zeros of the plant or the processing unit resulting in imperfect cancellation. As will be shown later, complete freedom in this regard is not always possible.

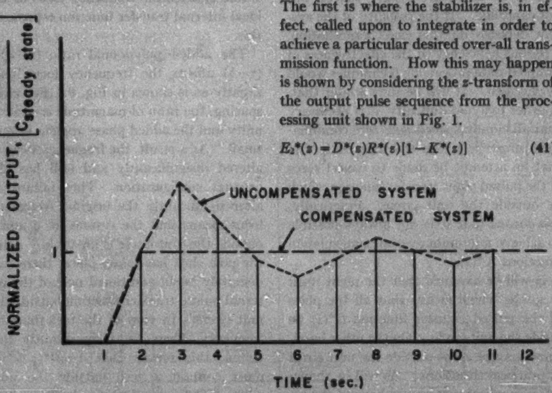


Fig. 7. Response of system used in illustrative example to step input

Straight-line interpolations not true envelope

1. Limitation Owing to Maximum Plant Input Variable

The fast response and finite settling time obtained with sampled-data processing technique is realized by "forcing" the plant input to any value necessary to realize the desired over-all response. In all practical plants, however, there will be an upper limit to the magnitude of the forcing function which the plant can accept. Thus, to prevent saturation and excessive forcing, the condition must be inserted in the design that for a given test input the error pulse sequence never exceed a prescribed magnitude. There are two situations which come into play to prevent this condition from being met. The first is where the stabilizer is, in effect, called upon to integrate in order to achieve a particular desired over-all transmission function. How this may happen is shown by considering the z-transform of the output pulse sequence from the processing unit shown in Fig. 1.

$$E_2^*(z)=D^*(z)R^*(z)[1-K^*(z)] \quad (41)$$

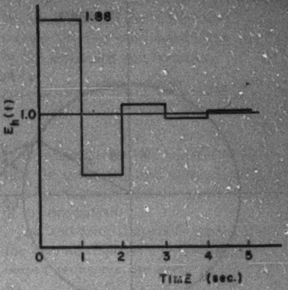


Fig. 8. Error signal for compensated system used in illustrative example

It has been shown in equations 13 and 14 that to close out the control error, $[1-K^*(z)]$ must contain a number of zeros at the point $z=1$ equal to the number of poles contained in the z-transform of the input, $R^*(z)$, resulting in a complete cancellation. Thus, should there be poles of $D^*(z)$ at $z=1$ they will not be cancelled. Any poles of order higher than the first will result in a pulse sequence at the output of the processing unit which tends to rise toward infinity and thereby overload the plant. If this is to be prevented, the rule may be stated that the sampled-data processing unit transfer function should not be required to contain poles at $z=1$ whose order is higher than the first.

Assuming that this condition is met, there is still the problem of preventing the pulse sequence applied to the plant from exceeding some prescribed magnitude during the transient following the application of some test function. If the pulse sequence at the output of the processing unit does so, it means that the desired over-all response $K^*(z)$ cannot be achieved with the available plant capability and that a modified response which most closely approximates the desired one must be sought. This is done by making note of the relation

$$K^*(z)=\frac{C^*(z)}{R^*(z)}=\frac{E_2^*(z)G^*(z)}{R^*(z)} \quad (42)$$

from which the equality results

$$K^*(z)R^*(z)=E_2^*(z)G^*(z) \quad (43)$$

Since the test function $R^*(z)$ and the plant transfer function $G^*(z)$ are assumed to remain fixed, the problem of specifying the closest feasible over-all transmission function $K^*(z)$ reduces to that of adjusting it so that no pulse in the error sequence $E_2^*(z)$ will exceed the prescribed magnitude limit. It is desirable that $K^*(z)$ be as short a polynomial as possible

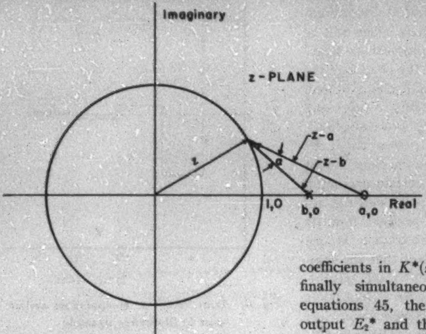


Fig. 9. Amplitude and phase effects of closely spaced pole and zero

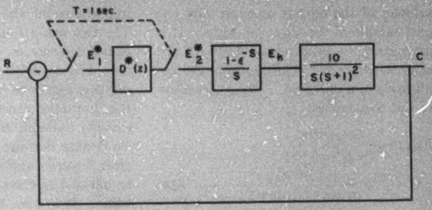


Fig. 10. System used in illustrative example

so that the settling time is minimal. It is required that the coefficients of this polynomial satisfy the restrictions imposed by equations 16 and 18.

Expressing the transforms in equation 43 in their series form, there results the following

$$[r_0+r_1z^{-1}+\ldots][k_mz^{-m}+k_{m+1}z^{-(m+1)}+\ldots k_pz^{-p}]=[e_0+e_1z^{-1}+\ldots]\times[g_mz^{-m}+g_{m+1}z^{-(m+1)}+\ldots] \qquad (44)$$

Multiplying this sequence through and collecting the coefficients of like orders of z, there results a number of equalities

$$\begin{aligned}&r_0k_m=e_0g_m\\&r_0k_{m+1}+r_1k_m=e_0g_{m+1}+e_1g_m\\&r_0k_{m+2}+r_1k_{m+1}+r_2k_m=e_0g_{m+2}+\\&\qquad e_1g_{m+1}+e_2g_m,\ \text{etc.}\end{aligned} \qquad (45)$$

Expressed in this form, and taking into account the relations of equations 16 and 18, it is possible to insert both the maximum allowable values of the error sequence resulting from a test input and the conditions on the coefficients of $K^*(z)$. For instance, taking the first relation in equation 45, r_0 is the pulse of the input function at the origin of time, g_m is the mth ordinate of the plant transfer function and e_0 is the error magnitude at the 0th sampling instant. The coefficient k_m is to be determined. If it is set at a value specified by a prototype function of Table I with a view toward realizing the fastest possible rise time, it may happen that e_0 exceeds the maximum magnitude permitted in the design. In this event, e_0 is set at the maximum value and k_m is computed. The same procedures are repeated in rote with the subsequent equations 45, always taking into account the conditions on the coefficients of $K^*(z)$ until a minimum number of coefficients in $K^*(z)$ are obtained which finally simultaneously satisfy all the equations 45, the maximum allowable output E_2^* and the conditions imposed on the coefficients. The procedure is relatively straightforward when numerical values are used, though it appears complicated when stated in general terms. The net effect on the over-all response function with this limitation included is to increase the finite settling time of the system as is to be expected. Considerations like these point the way to multiple state systems where the processor program is adjusted to match the magnitude of the error.

2. LIMITATIONS DUE TO IMPERFECT POLE-ZERO CANCELLATIONS

The desired internal transfer functions previously considered have been synthesized by cancelling out the undesired poles or zeros of the pulsed transfer function of the plant, $G^*(z)$, and replacing them with poles and zeros of a desired function $D^*(z)G^*(z)$. While perfect cancellation is ideally possible, this is not true in practice and the actual system will have poles and zeros which are closely spaced but not cancelled. If the stability of the system depends on this exact cancellation, a serious deficiency in the effectiveness of sampled-data processing techniques would be revealed. It will be shown later that, in order to avoid difficulty, it is necessary that all closely spaced pole-zero combinations must lie inside the unit circle and that no attempt be made to cancel zeros of the pulsed plant transfer function which lie outside the unit circle. Practically, this means that it is not always possible to achieve minimum over-all transmission functions $K^*(z)$.

It will be assumed that the plant itself is stable, which means that all the poles of the pulsed transfer function $G^*(z)$ lie inside the unit circle. On the other hand, there may be zeros outside the unit circle in practical situations. As will be shown later, practical system synthesis of the internal pulsed transfer function will not permit the cancellation of such zeros of $G^*(z)$ and these will therefore have to be included among the zeros of the over-all transfer function, $K^*(z)$.

To determine the effect of closely spaced pole-zero combinations which would result from imperfect cancellation use will be made of the frequency locus of the internal pulsed transfer function $D^*(z)G^*(z)$. The actual pulsed transfer function will be replaced by an ideal one in which perfect pole-zero cancellation is assumed to have taken place preceded by a closely spaced pole-zero pair or pairs. Some reflection will show that this combination is identical to a transfer function in which imperfect cancellation has taken place. This representation would be formulated for the case of a single pole-zero pair as

$$D^*(z)G^*(z)\big|_{\text{actual}}=\frac{z-a}{z-b}D^*(z)G^*(z) \qquad (46)$$

where a and b are respectively the zero and pole locations of the actual transfer function which do not exactly coincide and $D^*(z)G^*(z)$ is the ideal transfer function. Obviously, since the ideal over-all transmission function $K^*(z)$ is that of a stable system, the frequency locus of the ideal internal transfer function represents this.

The added polynomial ratio $(z-a)/(z-b)$ affects the frequency locus only slightly as is shown in Fig. 9. For close spacing, the ratio of magnitude is close to unity and the added phase angle α is very small. As a result, the frequency locus is altered insignificantly and still has the original configuration. This failure to alter significantly the original frequency locus means that the system for a and b outside the unit circle is unstable. After the pole shift has taken place, there will definitely be an additional pole of the internal pulsed transfer function outside the unit circle. In view of the fact that the frequency locus is insignificantly affected, this means that $1+D^*(z)G^*(z)$ must contain a zero outside the unit circle. Thus, a failure to obtain perfect cancellation poles or zeros of $G^*(z)$ outside the unit circle by $D^*(z)$ always results in an unstable system. Some consideration will show that this is not true for pole or zero cancellations inside the unit circle.

Summarizing the limitations on the design of the sampled-data processing unit

1. The unit should not be designed to integrate when a specified test function is applied since this would result in an indefinitely rising input to the plant.
2. The constants must be limited in a manner which does not exceed the maximum admissible input to the plant for the given test input function.
3. No zeros of the pulsed plant transfer function $G^*(z)$ lying outside of the unit circle can be cancelled by the processing unit.

None of these restrictions is particularly harmful to the system design. Observing them merely means that not all systems can be made to have over-all transfer functions which are minimum polynomials. The settling time in response to transient inputs may exceed the minimum and the overshoots may not be held to zero. On the other hand, even with these restrictions, the resulting systems have finite settling times at least in so far as values at sampling instants are concerned.

ILLUSTRATIVE EXAMPLE

To illustrate the design procedures discussed in the previous sections, an example of the design of the program for a sampled-data processing unit will be given. The system to be designed is shown in Fig. 10. It will be observed that this system is unstable even if it were continuous and will be even more so if it is error-sampled as shown. The pulsed plant transfer function, $G^*(z)$ is calculated in the usual manner resulting in

$$G^*(z)=\frac{(1+2.34z^{-1})(1+0.16z^{-1})(z^{-1})}{(1-z^{-1})(1-0.368z^{-1})^2} \qquad (47)$$

It is observed that $G^*(z)$ has a zero at -2.34, which cannot be cancelled by the processing unit owing to the restrictions mentioned previously. Thus, this zero must also appear as a zero in any over-all transmission function $K^*(z)$ selected

$$K^*(z)=(1+2.34z^{-1})F^*(z) \qquad (48)$$

Now if it is specified that the system be capable of following a ramp input function without error, there are the two conditions set forth in equations 20 and 21 which must be satisfied. This means that $F^*(z)$ must contain at least two terms with associated unspecified coefficients. Thus

$$F^*(z)=f_1z^{-1}+f_2z^{-2} \qquad (49)$$

and $K^*(z)$ becomes

$$K^*(z)=f_1z^{-1}+(2.34f_1+f_2)z^{-2}+2.34f_2z^{-3} \qquad (50)$$

This represents the shortest possible settling time obtainable for this plant which is, in this case, three sample periods. Applying next the conditions on the coefficients imposed by equations 20 and 21

$$\left.\begin{aligned}&3.34f_1+3.34f_2=1\\&f_1+2(2.34f_1+f_2)+3(2.34f_2)=0\end{aligned}\right\} \qquad (51)$$

from which the coefficients are 0.81 and -0.51 for f_1 and f_2 respectively. The overall prototype transmission function which can be realized is thus

$$K^*(z)=0.81z^{-1}+1.38z^{-2}-1.19z^{-3} \qquad (52)$$

The internal transfer function required to obtain this response is given by

$$D^*(z)G^*(z)=\frac{K^*(z)}{1-K^*(z)} \qquad (53)$$

Substituting $K^*(z)$ from equation 52

$$D^*(z)G^*(z)=\frac{(1+2.34z^{-1})(0.81z^{-1}-0.51z^{-2})}{(1-z^{-1})^2(1+1.19z^{-1})} \qquad (54)$$

Dividing through by $G^*(z)$, the pulsed transfer function which must be generated by the processing unit becomes

$$D^*(z)=\frac{(1-0.368z^{-1})^2(0.81-0.51z^{-1})}{(1-z^{-1})(1+0.162z^{-1})(1+1.19z^{-1})} \qquad (55)$$

It is evident from this transfer function that the processing unit must store three past input samples and three past samples of its own output.

If a unit ramp is applied to the system, the over-all response and processing unit output sequences are given respectively by

$$C^*(z)=\frac{D^*(z)G^*(z)}{1+D^*(z)G^*(z)}R^*(z) \qquad (56)$$

$$E_2^*(z)=\frac{D^*(z)}{1+D^*(z)G^*(z)}R^*(z) \qquad (57)$$

Substituting the z-transform of the unit ramp input $Tz^{-1}/(1-z^{-1})^2$, the respective sequences are obtained in the usual manner. The output pulse sequence and the output of the processing unit as it would appear after being clamped are shown in Fig. 11. The processing unit delivers an output which is a damped oscillation designed to result in a deadbeat performance for the over-all system after three sample periods. A practical note as to the relative lack of complexity of this unit is noted in the fact that a total of only six past samples need be stored.

Implementation of Sampled-Data Processing Units

Previous discussions have been devoted to analyses and theory which lead to a specification of the desired pulsed transfer function $D^*(z)$, which describes the sampled-data processing unit performance. Essentially this function specifies a pro-

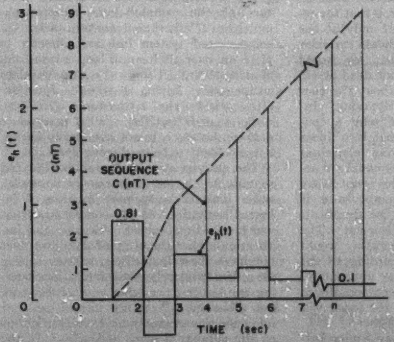


Fig. 11. Output and error pulse sequence for illustrative example

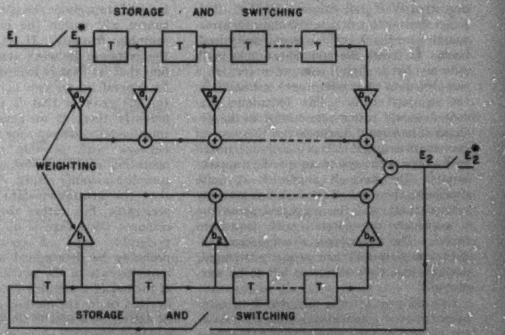


Fig. 12. Block diagram for data-processing unit

gram of weighting past samples of the input and output sequences and generating a pulse whose magnitude fulfills the requirements of the system. This section will concern itself with some considerations of implementation of these programs though not with circuit design and engineering considerations leading to particular choices of equipment.

In implementing the programs, it is noted that three basic operations are required. The first is that of data storage. Past samples must be stored in order that they may be weighted in the computation of the present output sample. This storage function may be implemented either with analogue or digital devices. For instance, if digital methods are used, the stored values would be retained in coded form on registers, magnetic tapes, or drums. On the other hand, if analogue methods are used, samples would be stored in the form of some analogue quantity such as a voltage, current, shaft rotation, mechanical, or volumetric displacement. The method which is chosen depends upon factors such as time scale, cost, accuracy, form of original data, and the data transmission devices to be employed.

The second operation involved arithmetic manipulations including multiplication or weighting of various samples by constants, addition and subtraction. These are all standard operations in computing devices whether they be implemented analogue-wise or digitally. Finally, there is a switching operation in which the various weights are applied sequentially at each sampling instant to the data samples in the manner prescribed by the program.

The form of the pulsed transfer function for the processing unit $D^*(z)$ can always be given by

$$D^*(z)=\frac{a_0+a_1z^{-1}+a_2z^{-2}+\ldots a_nz^{-n}}{1+b_1z^{-1}+b_2z^{-2}+\ldots b_mz^{-m}} \qquad (58)$$

This implies that the relation between the input and the output pulse sequence must be

$$\begin{aligned}E_2(nT)=a_0E_1(nT)+a_1E_1(n-1)T+\ldots\\ -b_1E_2(n-1)T-b_2E_2(n-2)T-\ldots \quad (59)\end{aligned}$$

A literal block diagram representing this particular operation schematically is shown in Fig. 12. The rectangular blocks in this figure indicate time delays which represent the storage and switching operation. The weights, a_0, $a_1, \ldots a_n$ and b_1, $b_2 \ldots b_m$, represent the weighting sequence which is applied to the various samples.

The exact form of implementation which is used in any particular application depends on the various factors mentioned previously as well as the investment represented by the plant. For instance, in a plant of large magnitude in this respect, where many independent control systems may be required, it is possible to take advantage of the high computing capacity of a digital system which can then be multiplexed among a number of processes. In isolated installations where low computing capacity is required, analogue methods may be more economical. The exact form of implementation cannot be specified until the exact process and its environment is specified. In any case, it is not beyond the realm of straightforward engineering practice to obtain a satisfactory implementation of the programs required for process control. As a matter of fact, as mentioned in the early part of this paper, the flexibility of design and control possible with sampled-data processing units may make their employment in continuous systems desirable even in cases where the sampled-data system is not imposed by circumstances.

References

1. A New Approach to the Design of Pulse Monitored Servo Systems, **A. Porter, F. Stoneman.** *Journal*, Institution of Electrical Engineers, London, England, vol. 97, pt. II, 1950, pp. 597–610.

2. Sampled-Data Control Systems Studied Through Comparison of Sampling With Amplitude Modulation, **William K. Linvill.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1779–88.

3. The Analysis of Sampled-Data Systems, **J. R. Ragazzini, L. A. Zadeh.** *AIEE Transactions*, vol. 71, pt. II, 1952, pp. 225–34.

4. A General Theory of Sampling Servosystems, **D. F. Lawden.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 98, pt. IV, Oct. 1951, pp. 31, 36.

5. Analysis and Synthesis of Sampled-Data Control Systems, **E. I. Jury.** *AIEE Transactions*, vol. 73, pt. I, Sept. 1954, pp. 332–46.

6. The Pulse Transfer Function and Its Application to Sampling Servo Systems, **R. H. Parker.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 99, pt. IV, Dec. 1952, pp. 302–17.

7. Treatment of Digital Control Systems and Numerical Processes in the Frequency Domain, **J. M. Salzer.** *D.Sc. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951.

8. Analysis of Control Systems Involving Digital Computers, **W. K. Linvill, J. M. Salzer.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 41, no. 7, July 1953, pp. 901–06.

9. Numerical Methods in Engineering, (book) **M. G. Salvadori, M. J. Baron.** Prentice-Hall, Inc., New York, N. Y., 1952, pp. 51–52.

Discussion

John M. Salzer (Magnavox Research Laboratories, West Los Angeles, Calif.): This paper deals with a design technique for error-sampled-feedback systems, using unity feedback. It treats the compensation of such systems by a digital computer (or by a sampled-data processing unit) operating on the sampled error. The technique presented should prove very useful to the designer if he should decide to use this type of compensator and if he is primarily interested in response of the system to specific transient inputs. The over-all discussion is quite thorough, the method presented is strongly systematized, and the suggested procedure is copiously documented with transform tables. The advantages of using sampled-data compensation are properly stressed; nevertheless, their review is felt to be warranted.

One advantage of a sampled-data compensator over a continuous-data compensator stressed in the paper is that a deadbeat response to a step or ramp input can be achieved. The expression "deadbeat" means that after a fixed and finite time interval the response is exactly as the input. What needs additional emphasis, although implicit throughout the paper, is that the response will match the input only at the sampling instants. If we should consider the response between samples, we would find that its beat is indeed not dead at all. In general all we can say about the continuous error is that it is decaying. Inherently there is no reason why a continuous-data compensator (that is, a linear network with possible active elements) operating on the sampled error could not accomplish a similar result. This is not to say that the continuous-data compensator is preferable, but rather that the deadbeat-response advantage is not clear-cut. The preferable type of compensator would probably be determined according to the requirements of the particular application. If the use of a sampled-data compensator is decided upon, the method of the paper will be extremely useful to the designer.

The advantage of employing a sampled-data compensator in a sampled-data system is that the exact analytical work is vastly simpler than it would be with a continuous-data unit. As pointed out by the authors, this is so because $D^*(z)$, the transfer function of the sampled-data compensator, multiplies $G^*(z)$, the z-transform of the uncompensated system function, directly to yield an over-all forward-loop z-transform of $D^*(z)G^*(z)$. Using a continuous-data compensator, having a transfer function $H(s)$, yields the z-transform $(HG)^*(z)$, rather than $H^*(z)G^*(z)$. What is easier to analyze, however, is not necessarily better, but it is likely to be used more often.

The authors state that stabilization and shaping, by the use of continuous networks, suffer limitations in part because of inherent limitations of the control functions which can be obtained. Clearly, the implication is that the sampled-data unit does not have these limitations, whatever they are. Unfortunately, these limitations binding one but not the other kind of filter are not discussed. It is my impression that reference was probably made to restrictions on the number of zeros and poles. In the practical case, the transfer function of a continuous filter can have no more zeros than poles. No such restriction hampers the sampled-data filter. As a matter of fact, $D^*(z)$ can be void of poles altogether, in which case it is a polynomial in z^{-1}, rather than a rational function. This freedom could easily turn out an advantage in system compensation. Although the method of the paper can conceivably lead to $D^*(z)$ having more zeros than poles, the illustrations do not, and in general this possibility is not exploited.

The response to arbitrary inputs of a system designed for step or ramp inputs is an important problem, which is not neglected by the authors. Indeed, this question is very well discussed (even though the derivations are curt) and neatly correlated with results of numerical analysis. The relation between transient and steady-state (sinusoidal) responses is at least heuristically explained.

A comment in the paper regarding a limitation on $D^*(z)$ raises some questions. To prevent the plant from overloading $D^*(z)$ may have at most one pole at $z=1$. Now

$$1-K^*(z)=\frac{1}{1+D^*(z)G^*(z)}$$

or after writing $D^*(z)G^*(z)$ in terms of its numerator and denominator polynomials

$$1-K^*(z)=\frac{1}{1+P^*(z)/Q^*(z)}=\frac{Q^*(z)}{P^*(z)+Q^*(z)}$$

However, equation 14 of the paper requires that $1-K^*(z)$ and therefore $Q^*(z)$ has the factor $[1-z^{-1}]^m$ or that $Q^*(z)$ has m zeros at $z=1$, where m is the order of the test input function ($m=1$ for step, $m=2$ for ramp, etc.). Since $Q^*(z)$ is the denominator of $D^*(z)G^*(z)$, the latter must have m poles at $z=1$. But the limitation under discussion allots at most one such pole to $D^*(z)$ and consequently imposes the following requirement on the plant $G^*(z)$: in order to have a zero steady-state sampled error to an $m-1$ degree polynomial input the system must have a $G^*(z)$ with at least $m-1$ poles at $z=1$. How serious is such a limitation in practical cases?

The method of adjusting the parameters of $D^*(z)$, so as to keep excessive signals from reaching the plant, demonstrates the use of straightforward numerical calculations. The usefulness of simple numerical approaches to other detail questions in sampled-data systems is indicated.

R. E. Kalman (E. I. du Pont de Nemours & Company, Wilmington, Del.): The authors should be commended for their stimulating treatment of the compensation of linear control systems by means of discrete filters (sampled-data processors). Their method constitutes time-domain synthesis, and follows the current trend in the feedback control field which emphasizes time-domain concepts such as the root locus and impulse response in contrast to the older attack which utilizes the frequency response. The important advantage of the time-domain approach is that it represents the transient behavior of a system in direct physical terms; this constitutes a much broader viewpoint than the frequency-response representation (which, strictly speaking, is applicable only to linear or mildly nonlinear systems), and frequently gives deeper insight into the problem.

In illustrating the power of the time-domain philosophy, I should like to outline an alternate method for determining the discrete compensating filter $D^*(z)$ when the plant (output element) $H(s)$ and the form of the desired transient response are specified. This approach, in a strikingly simple fashion, enables one to base the design directly on the properties of $H(s)$ and at the same time provides a very elegant qualitative, as well as quantitative, understanding of the system dynamics. In particular, it will be shown how the selection of the sampling period T becomes an integral part of the design procedure. Moreover, the reasoning to be used here is identical to that encountered in conjunction with the so-called "optimum relay servos"[1-4] where the problem is, loosely speaking, to achieve the fastest possible transient response in the presence of sharp saturation. Since the optimum relay servo is one of the very few strongly nonlinear systems which are fairly well understood at present, it is hoped that this note will aid in clarifying the often vague distinctions between linear and nonlinear dynamic systems via simple physical principles.

This approach will be explained with the aid of an example. Let us assume that the plant has two real poles, i.e., $H(s)=ab/(s+a)(s+b)$. We wish to examine the action of a position control system in response to a unit step input. It is clear that the input $e_h(t)$ to the plant $H(s)$ is unity in the steady state (cf. Fig. 1) if we specify a zero position error system. To achieve fast response, it is natural first to accelerate the plant by applying a forcing signal much larger than unity; then after the transient error has sufficiently diminished, one should apply a strong decelerating signal (smaller than unity). To keep the analogy to the optimum relay case as close as possible, we now stipulate, for the assumed $H(s)$, that the transient response to a unit step should have no overshoot and should persist for the duration of two sampling periods such that at $t=2T$ the error is reduced exactly to zero, and stays zero not only at the sampling point but also in between. Under such assumptions, it is easy to see that the qualitative shapes of $e_h(t)$ and $c(t)$ must be as sketched in Fig. 13, where the dotted lines indicate the shape of the transient when $e_h(t)$ is maintained at a constant level. Stated differently, as a consequence of these assumptions $e_h(t)$ will contain a damped transient term of finite duration, whereas with the authors' method the transient will have infinite duration (cf. Figs. 8 and 11). In the latter case, the continuous output $c(t)$ will always have an exponentially decaying ripple (i.e., overshoots between sampling intervals, similar to an oscillatory transient in a system without sampling) which is frequently objectionable. By the foregoing definition of the optimum response, the ripple is automatically eliminated if $D^*(z)$ is determined by the method which follows.

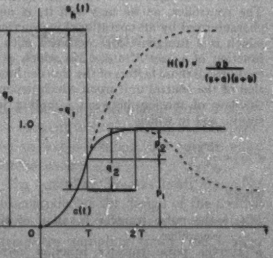


Fig. 13. Sketch of forcing signal $e_h(t)$ and output $c(t)$ in case of optimum (nonovershooting) transient response to a unit step

The conditions inherent in Fig. 13 may be readily stated in mathematical form. In fact we must have, in the terminology of Fig. 1, where $G^*(z)=[(1-z^{-1})H(s)/s]^*$

$$\frac{E_h^*(z)}{R^*(z)}=\frac{D^*(z)}{1+D^*(z)G^*(z)}=q_0+q_1z^{-1}+q_2z^{-2}=Q^*(z) \quad (60)$$

$$\frac{C^*(z)}{R^*(z)}=\frac{D^*(z)G^*(z)}{1+D^*(z)G^*(z)}=p_1z^{-1}+p_2z^{-2}=P^*(z)$$

and, because of steady-state considerations

$$\sum_{i=1}^{2}pi=1$$
$$\sum_{i=1}^{3}qi=1 \quad (61)$$

By comparing the two equations 60, we conclude that

$$G^*(z)=\frac{p_1z^{-1}+p_2z^{-2}}{q_0+q_1z^{-1}+q_2z^{-2}}=\frac{P^*(z)}{Q^*(z)} \quad (62)$$

Since $G^*(z)$ is a rational function in z^{-1}, the equality 62 can obviously be satisfied; moreover, as we may choose the gain in the system, which is included in $D^*(z)$, so as to set $|H(0)|=1$, conditions 61 may also be met by multiplying numerator and denominator of $G^*(z)$ by suitable scale factors such that the restriction $|G^*(1)|=1$ is preserved. Then it is easy to find $D^*(z)$ from either of conditions 60

$$D^*(z)=\frac{Q^*(z)}{1-G^*(z)Q^*(z)}=\frac{Q^*(z)}{1-P^*(z)} \quad (63)$$

The important result here is equation 62. It shows that to compute the various significant parameters in Fig. 13, one must merely find the z-transform of the hold circuit-plant combination $(1-z^{-1})H(s)/s$. The only undetermined factor is then the sampling period T. Since we are dealing with a linear system, the transient response could be speeded up arbitrarily by decreasing T. This, however, would require larger and larger initial signals into the plant $H(s)$, i.e., $q_0 \rightarrow \infty$. But q_0 naturally cannot be increased indefinitely owing to saturation or other types of power limitations, and so there will always be some ultimate bound on the rapidity of the transient response. In short, the selection of T is limited only by the available power levels in the system, but not by the time constants. Since the effects of power limitation are a fundamentally nonlinear problem they cannot be considered here, except to note that one could perhaps design nonlinear discrete filters to compensate for such nonlinear effects.

As a numerical illustration, consider $H(s)=3/(s+1)(s+3)$. If the sampling period is $T=1$, we have, multiplying numerator and denominator by suitable scale factors to satisfy equation 61

$$G^*(z)=\left[\frac{1-e^{-s}}{s}H(s)\right]^*$$
$$=\frac{0.786z^{-1}+0.214z^{-2}}{1.66-0.69z^{-1}+0.03z^{-2}} \qquad (64)$$

But if T is lowered to 0.2, i.e., speeding up the response by a factor of 5

$$G^*(z)=\left[\frac{1-e^{-0.2s}}{s}H(s)\right]^*$$
$$=\frac{0.574z^{-1}+0.426z^{-2}}{12.35-16.9z^{-1}+5.55z^{-2}} \qquad (65)$$

Bearing in mind the definition of the coefficients in the numerator and denominator of equations 64 and 65 as given by equations 60 and Fig. 13, it is seen that the ratio of control effort required when $T=0.2$ to that when $T=1$, i.e., $q_0(T=0.2)/q_0(T=1.0)=12.35/1.66=7.44$ is considerably larger than the improvement in response time. Thus q_0, evaluated for a given T, is a most important piece of quantitative design information. Other coefficients of equations 64 and 65 may be similarly interpreted on the basis of Fig. 13.

It may be verified by direct computation that $D^*(z)$ corresponding to equations 64 or 65 contains a pole at $z=1$, i.e., the compensating network acts as an integrator which is what one would expect from a servo with zero steady-state positional error.

Although further detailed discussion is not possible here, it should be pointed out that, in view of equation 62, the number of steps in $e_h(t)$ before reaching the steady-state value is determined by the degree of the polynominal $Q^*(z)$. This is in agreement with the results of Bogner[5] and the writer[6] in connection with high-order optimum relay servos.

Thus both the optimum relay servo case (which is highly nonlinear) and the design of a discrete compensator (with linear but discontinuous output) are based on the same simple physical principle, which involves the question: In what way should steps be applied to a plant in order to achieve zero steady-state position error as quickly as possible, with a nonovershooting transient? In an optimum relay servo, the absolute level of the forcing signal is usually maintained constant (at forward or reverse saturation) and the time of application is varied; in a compensated sampled-data system the amplitude of the forcing signal is varied while the time is fixed by the sampling period.

REFERENCES

1. NONLINEAR TECHNIQUES FOR IMPROVING SERVO PERFORMANCE, **D. C. McDonald.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 6, 1950, pp. 400–21.

2. PHASE-PLANE APPROACH TO THE COMPENSATION OF SATURATING SERVOMECHANISMS, **A. M. Hopkin.** *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 631–39.

3. THE APPLICATION OF NONLINEAR TECHNIQUES TO SERVOMECHANISMS, **K. C. Mathews, R. C. Boe.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 8, 1952, pp. 10–21.

4. DESIGN CONSIDERATIONS OF A SATURATING SERVOMECHANISM, **P. E. Kendall, J. F. Marquardt.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 9, 1953, pp. 178–87.

5. AN INVESTIGATION OF THE SWITCHING CRITERIA FOR HIGHER ORDER CONTACTOR SERVOMECHANISMS, **I. Bogner, L. F. Kazda.** *AIEE Transactions*, vol. 73, pt. II, July 1954, pp. 118–27.

6. SUMMARY OF WORK ON RELAY-TYPE NONLINEAR SERVOMECHANISMS, **R. E. Kalman.** *MIT DIC Project 6897, Internal Memorandum 43*, Massachusetts Institute of Technology, Cambridge, Mass., July 1953.

R. H. Barker (Joint Services Staff College, Latimer, Chesham, Buckinghamshire, England): This paper does not describe any fundamentally new method of synthesising an error-sampled control or servo system. The over-all transmission function $K^*(z)$ is selected to follow a test function of the unit step, ramp, or acceleration type without steady-state error by including $1-z^{-1}$ to the appropriate power as a factor of $1-K^*(z)$ and to fulfill the requirement of finite settling time. These conditions were both considered in reference 6, and in fact finite settling time can be obtained only if the characteristic equation is of the form $z^N=0$.

However, that a minimum settling time is desirable is very debatable. If the system is intended to follow high-order derivative functions with zero steady-state error, it will have a transient response containing large overshoots and considerable magnification of random errors. Some degree of smoothing is usually essential, and the use of characteristic equations of standard form analogous to those of Whiteley[1] is probably as good a method as any.

The conclusion that poles and zeros cannot be cancelled if they lie outside the unit circle is noteworthy. Any attempt to do this results in what was called high-frequency instability in reference 6.

The z- or sequence-transform method yields information only at the sampling instants. In Fig. 1 the hold system may generate a simple step function or it may be elaborated to impose a particular law of motion on the output between sampling instants. The processing unit then controls the output at these sampling instants. The intermediate sampler enables these two parts to be designed separately. If, however, the processing unit is an analogue device, considerable economy of equipment may result if it can incorporate also the hold system so that the intermediate sampler is omitted. The design is now more difficult since equation 4 no longer applies. The controller, as we may call it, is now characterized by an operational instruction, which is a function of z combined with a function of s, say $\theta^*(z)\phi(s)$, in which $\phi(s)$ contains all those factors of the transfer function of the control unit upon which depend the law of motion between sampling instants, and in which

$$\theta^*(z)=\frac{D^*(z)}{\phi^*(z)}$$

where

$D^*(z)=$ the pulse transfer function of the controller
$\phi^*(z)=$ the pulse transfer function corresponding to $\phi(s)$

To use the operational instruction the transfer function of the plant is multiplied by $\phi(s)$ and the pulse transfer function corresponding to the product is multiplied by $\theta^*(z)$. The result is used instead of $D^*(z)G^*(z)$ in equation 4.

To cover one or two details: 1. The sentence following equation 6 is loosely worded. If $\tau=0$ Fig. 3 shows that $g(t)$ begins to rise immediately and it is only the sampled value which is not other than zero until $t=T$. 2. On a point of nomenclature, the expression "pulse transfer function" rather than "pulsed transfer function" is generally accepted. The latter would imply that the function is somehow pulsed, whereas it actually relates to the transfer of pulses (or samples) from input to output.

REFERENCE

1. THEORY OF SERVO SYSTEMS WITH PARTICULAR REFERENCE TO STABILISATION, **A. L. Whiteley.** *Journal*, Institution of Electrical Engineers, London, England, vol. 93, pt. II, 1946, p. 361.

Arthur R. Bergen and **John R. Ragazzini**: We wish to thank the discussers for their interesting and constructive comments.

Dr. Salzer's discussion raises several important points. As he correctly points out, deadbeat response at sampling instants is not deadbeat in between samples at all. Our specifications of $K^*(z)$ should therefore be checked in every case by a ripple analysis. Methods for evaluating the ripple are available.[1,3] It is clear that a more sophisticated synthesis procedure would be desirable which would optimize the continuous response rather than the response only at sampling instants.

Our statement regarding the inherent limitations of continuous networks as compared to sampled-data units should be altered to read: "Stabilization and shaping, by the use of continuous networks, suffer limitations due, in part, to inherent limitations of the control functions which can be obtained with practical networks." We recognize that theoretically filters can be designed to accomplish the same stabilization as with sampled-data units.

Generally speaking, it is more difficult to stabilize sampled-data systems than continuous systems, and simple network compensation of the lag or lead type may not be effective. While more effective network arrangements are possible, these are more difficult to design, adjust, and maintain in adjustment. In addition, the values of components called for may not be practical. Therefore, in many cases a more practical solution is in terms of the sampled-data processing unit.

Dr. Salzer raises the question of how serious a practical limitation is the requirement that any poles of $D^*(z)$ at $z=1$ be simple. In discussing this question, it should be recalled that this restriction was imposed in order to secure a finite input to the plant. The basic engineering consideration is that of plant saturation; therefore, the restriction on $D^*(z)$ may not apply in many practical cases. For example, if the system input were a ramp from zero to infinite time, then it certainly would be necessary for the plant to have an integration to avoid saturation. However, if this ramp were applied for only a finite time, it might be possible to follow at sampling instants without a plant integration. The restriction on $D^*(z)$ regarding the number of poles at $z=1$ may therefore be an unnecessarily severe restriction brought about by applying a somewhat artificial test function.

Mr. Kalman's method of obtaining a transient response without ripple or overshoot is very interesting. A complete treatment of the material should be an important contribution to the field. For many applications the method seems entirely satisfactory as it stands. For some applications, however, the method as presently described does not give enough flexibility to the designer. For example, if T is fixed by considerations other than that of available power, then for any given hold and plant the design is uniquely determined and may in fact be unacceptable. However, it should be possible to gain more flexibility in design by specifying compensating networks in cascade with the plant. The difficulty here is in the shaping of the numerator polynomial of the pulsed transfer function of the hold, plant, and compensating network, in terms of the compensating network transfer function. Another difficulty arises if the coefficients of $P^*(z)$ are not all positive, in which case there may be undesirable overshoots, although the ripple is still eliminated.

We have long been acquainted with the outstanding work done by Mr. Barker and his British colleagues in the field of sampled-data systems. Reference 6 of the paper is a classic in the field and should be read by all who are interested in learning the fundamentals of sampled-data systems. However, in the light of the first paragraph of Mr. Barker's discussion, we would like to point out some basic limitations in his work which have been eliminated in this paper. First, it is true that reference 6 states the conditions for perfect follow of a step, ramp, or acceleration. Also, Mr. Barker has considered the problem of finite settling time which he achieves by forcing the characteristic equation to take the form of $z^N = 0$. This paper, however, goes well beyond the mere condition of finite settling time. Not only is the settling time made finite but also the shape of the transient response at sampling instants is controlled and subjected to the condition of maximum power available at the plant input. By working only with the characteristic equation, as is done in reference 6, the shape of the transient is not under the control of the designer and may not represent desirable performance.

We agree with Mr. Barker that minimum finite settling time is not always desirable as a general design objective. Perhaps by overstressing the idea of minimum settling time, we have implied its general desirability. This is not actually the case and, as Mr. Barker points out, minimum prototype functions may perform quite poorly when subjected to inputs with lower order derivatives than designed for. In design the type of input is taken into account in setting up the desired over-all transmission $K^*(z)$ to fit the requirements, and this may or may not be the minimum settling time prototype.

The "high-frequency instability" referred to by Mr. Barker in reference 6 bears no relation to the effect resulting from the imperfect cancellation of zeros of $G^*(z)$ outside the unit circle. The high-frequency instability he refers to is intersample instability which we called "ripple." Imperfect cancellation of a zero of the transfer function lying outside the unit circle results in an over-all instability of the system. Stated differently, the output would become unbounded both at sampling and intersample instants by such an imperfect cancellation.

We cannot agree with the remarks concerning the lack of validity of equation 4 when applied to a sampled-data processing unit containing internal clamp circuits. Most digital and analogue implementations (including one being constructed by the authors) contain internal clamp circuits, and it would be a serious limitation indeed if the theory were not applicable. Referring to Fig. 1, the hold system is combined with the plant only for purposes of analysis and computation. Actually, the dotted line enclosing the part of the system described $G^*(z)$ need merely be redrawn so that the hold system is part of the processing unit if such be the desirable condition. The mathematics is the same whether the hold system is considered part of the plant or part of the processing unit. If the latter, the output of the processing unit would be E_h rather than E_2^*.

The nomenclature in this field is a serious problem which can only be stabilized by actions of appropriate standards committees. It happens that neither "pulse transfer function" nor "pulsed transfer function" is yet generally accepted. In support of this view, we refer the reader to an excellent report by Mr. Barker,[3] where he used "pulsed" throughout and also to the more recent discussion by one of his colleagues,[4] where "pulsed" was used again. As a matter of fact, given a choice, we would prefer "pulsed" since it implies a statement "pulsed filter or system transfer function" where the qualifying words "filter," "system," etc. are omitted. It should be pointed out that any transfer function is the transform of the response of a linear system to a short pulse or impulse and that the term "pulse transfer function" has its semantic limitations too.

REFERENCES

1. EXTENSION OF CONVENTIONAL TECHNIQUES TO THE DESIGN OF SAMPLED-DATA SYSTEM, W. K. Linvill, R. W. Sittler. *Convention Record*, Institute of Radio Engineers, New York, N. Y., 1953, pp. 99–104.

2. ANALYSIS OF ERRORS IN SAMPLED-DATA FEEDBACK SYSTEMS, J. Sklansky, J. R. Ragazzini. *Electronics Research Laboratories Technical Report T-3/B*, Columbia University, New York, N. Y., Feb. 1954.

3. THE THEORY OF PULSE MONITORED SERVOMECHANISMS AND THEIR USE FOR PREDICTION, R. H. Barker. *SRDE Report No. 1046*, Sig. Res. and Dev. Establishment, Christchurch, Hants, England, Nov. 1950.

4. *Discussion* by Prof. C. Holt Smith. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 101, part IV, no. 6, Feb. 1954, p. 162.

The Design of Sampled-Data Feedback Systems

GLADWYN V. LAGO
MEMBER AIEE

JOHN G. TRUXAL
ASSOCIATE MEMBER AIEE

◆

SAMPLED-DATA feedback control systems, or pulsed servomechanisms, are characterized by the distinguishing feature that at one or more points in the system signal information is in the form of a pulse train which can be represented as a sequence of numbers. One example in which the signal information inherently is represented by a sequence of numbers is provided by a closed-loop system containing a digital computer as one element. Sampled-data systems are also useful when the energy available from a sensitive element is so small that it is imperative to avoid loading the output more than necessary.

The sampling process both simplifies and complicates the design of the control system. Because the loop is closed only during the sampling times, the system operates most of the time open-loop, and evaluation of the response is a simple open-loop analysis problem. The intermittency of the closing of the loop introduces, however, stability problems nonexistent in the continuous system. The following discussion indicates certain basic ideas concerning the design of sampled-data systems, particularly the manner in which the simplification introduced by the sampling can be utilized in design.

The system to be considered is shown in Fig. 1. The distinguishing component is the sampler, which is represented as a switch alternately closed for T_1 seconds

---

Paper 54-283, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 5, 1954; made available for printing April 23, 1954.

GLADWYN V. LAGO is with the University of Missouri, Columbia, Mo., but is on leave of absence to complete work toward the Ph.D. degree at Purdue University on a Radio Corporation of America fellowship in electronics, as administered by the National Research Council. JOHN G. TRUXAL is with Purdue University, Lafayette, Ind.

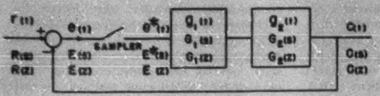


Fig. 1. A sampled-data feedback system

and opened for T_2 seconds. The sampling period T, assumed constant in this paper, is defined by the equation

$$T_1+T_2=T \tag{1}$$

A sampled-data system implies, by definition, that T_1 is small compared with both T_2 (or T) and the time constants of the remainder of the system.

For the purposes of illustration throughout this paper, only the single-loop unity-feedback configuration of Fig. 1 is considered. Here $G_2(s)$ is the transfer function of the controller elements and $G_1(s)$ describes the compensation network or control system. The product of $G_1(s)$ and $G_2(s)$ is denoted by $G(s)$. Furthermore, the primary emphasis of the following discussion is placed on the design to meet specifications phrased in terms of the desired characteristics of the step-function response, although the methods are readily extended to the consideration of other aperiodic inputs.

The continuous actuating signal is $e(t)$, while the actuating signal in sampled form is $e^*(t)$. The $e^*(t)$ has the form of a train of pulses with a width T_1 and a height proportional to the magnitude of $e(t)$ at the sampling instant. As long as the assumptions of the smallness of T_1 are valid, the analysis can be simplified by approximating the pulse train by a train of unit impulses with areas representing the sample amplitudes of $e(t)$. Substitution of the impulse train for the actual pulse train, with a correspondence of area to amplitude, tacitly assumes that T_1 equals unity. When the gain K of the open-loop part of the system is set on the basis of this assumption, the actual gain used in the physical system must be K/T_1.

Simplicity of System Response

By its very nature a sampled-data system should be more easily analyzed than the corresponding continuous system. Indeed, some of the most powerful methods for the analysis of continuous systems are based on sampled approximations of the continuous functions. The greater simplicity of sampled-data feedback systems results from the open-loop nature of the operation at all times except the sampling instants. The response of the over-all system of Fig. 1 is simply the sum of weighted impulse responses of g.

If the system shown in Fig. 1 is at rest and an arbitrary input function $r(t)$ is applied, the g network remains unexcited until the first sampling instant t_0. As a result of the first sampling, g is excited by an impulse of area

$$e(t_0)=r(t_0)-c(t_0) \tag{2}$$

In the interval $t_0<t<t_0+T$ the output $c(t)$ is simply the constant $e(t_0)$ multiplied by $g(t)$, the impulse response of g. At the second sampling period another impulse of area $e(t_1)$ is applied to g. Since g is assumed linear superposition holds and the output $c(t)$ is the sum of the two impulse responses.

If this procedure is continued, the output $c(t)$ can be calculated through any desired number of sampling periods. The output during the nth interval is simply the sum of n weighted, delayed impulse responses, in contrast to the continuous system where the output is the convolution of the actuating signal and $g(t)$. If an aperiodic function such as a unit step or a unit ramp function is applied to the system, the nature of $c(t)$ can often be determined by following the output through only a few sampling periods. The main purpose of this paper is to indicate the manner in which this physical view of system performance can be utilized in design; indeed, a wide variety of design problems can be solved with only this physical reasoning and without the introduction of more powerful mathematical methods.

Of the more elaborate mathematical tools available for the analysis and design of sampled-data systems the most promising appears to be the z-transformation.[1] The main limitation of the z-transform is that the response is determined only at the sampling instants. Since there exists an infinite number of sampled-data systems which have the same response values at the sampling instants, other techniques must be used to determine the nature of the complete response. The significance of this ambiguity is illustrated by the four systems shown in Fig. 2. The step-function response of each of the four systems assumes exactly the same set of values at the sampling instants, but the curves between sampling instants are remarkably different, as shown in Fig. 3. Since all four responses go through the same points at the sampling instants, the four associated z-transforms are identical. This uncertainty introduced in the use of the z-transform is added to the ambiguity introduced when the pulse train is replaced by an impulse train: an ambiguity which arises because a true impulse is also represented as an impulse in the approximation. For example, a fifth system can be added to Figs. 2 and 3; i.e., a system which possesses a step-function response which is entirely an impulse train.

A large part of this limitation of the z-transformation method is removed by the techniques set forth in Appendix I. These methods allow determination of the value of the output at any desired submultiple of the sampling period. Actually, this extended z-transformation approach is quite similar to the impulse response approach proposed in this paper; the main difference is primarily a matter of emphasis. If the extended z-transformation method is used as a basis for design, the designer is apt to become so enamored with the mathematics that all problems are worked with transformation methods. In the design of sampled-data systems it is important that the basic simplicity of system operation be kept firmly in mind. This can be accomplished if the basic approach to design is through the summation of the weighted

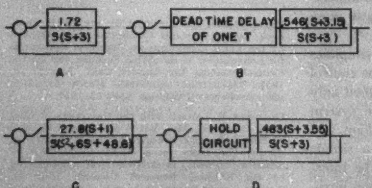


Fig. 2 (left). Four systems with the same z-transform and a T of 1 second

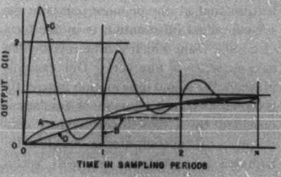
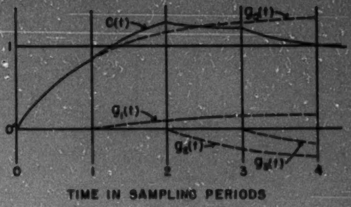


Fig. 3 (right). Step-function responses of the systems of Fig. 2

impulse responses of g. The z-transform, the theory of difference equations, and other mathematical methods are introduced to extend the designer's insight, to suggest system designs which may not be intuitively apparent, and to facilitate those aspects of analysis where intuitive and physical reasoning breaks down.

Fig. 4. Response of simple system



Impulse Response Approach

ANALYSIS

The basic characteristics and advantages of the impulse response approach to system design are conveniently described by considering first the analysis problem. In particular, the following simple example illustrates that analysis based on the impulse response of the open-loop system is capable of presenting essentially complete information about the system output. If a plot of the continuous output is not required, the output at regular intervals can be determined. The work required in the impulse response approach depends only on the amount of information desired.

The characteristics of the analysis are apparent from the simple example of Fig. 1 with

$$G(s)=\frac{13.4}{s(s+10)} \quad \text{and} \quad T=0.1 \text{ second} \tag{3}$$

The response of g to a unit impulse is

$$g(t)=1.34(1-\epsilon^{-10t}) \tag{4}$$

It is assumed that the system is at rest when the unit step function is applied as the input $r(t)$.

The first sampling (assumed at $t=0$, just after the application of the step) occurs with the actuating signal equal to unity. Thus at $t=0$ a unit impulse is applied to g which responds in accordance with equation 4 until the next sampling instant. This response is shown as $g_0(t)$ in Table I and Fig. 4. At the end of one sampling period the output has reached a value of 0.846, the error e_1 is 0.154, and an impulse of area 0.154 is applied to g. The corresponding component of the response is shown as $g_1(t)$ in Table I and Fig. 4. In Table I the $g_1(t)$ column is simply the $g_0(t)$ column multiplied by e_1 and shifted down by an amount corresponding to one sampling period. The output $c(t)$ between T and $2T$ seconds is

$$c(t)=g_0(t)+g_1(t) \tag{5}$$

and is also shown in Table I and Fig. 4. The remainder of the $c(t)$ curve is constructed in a similar manner, with the procedure followed for as many sampling periods as desired. Examination of the nature of $g_0(t)$ and the $c(t)$ curve of Fig. 4 shows that the output is settling down rapidly.

It should be pointed out here that the entries in Table I are simply the coefficients which are involved if the z-transform analysis is used with a subharmonic period of $T/5$. See Appendix I. The advantage of viewing the analysis as the superposition of impulse responses essentially lies in the possibility of utilizing the curve summation indicated in Fig. 4. In practice it is rarely necessary to calculate accurately more than the response at the actual sampling points. A knowledge of the graphical nature of $g(t)$ enables the designer to sketch the approximate shape of the response between sampling points without elaborate calculations.

Table I. Response of Simple System

t	$e_0=1$ $g_0(t)$	$e_1=0.154$ $g_1(t)$	$e_2=-0.29$ $g_2(t)$	$e_3=-0.204$ $g_3(t)$	c(t)
0	0				0
0.02	0.242				0.242
0.04	0.442				0.442
0.06	0.605				0.605
0.08	0.738				0.738
0.10	0.846	0			0.846
0.12	0.935	0.037			0.972
0.14	1.01	0.068			1.078
0.16	1.07	0.093			1.163
0.18	1.12	0.114			1.234
0.20	1.16	0.130	0		1.29
0.22	1.19	0.144	−0.07		1.264
0.24	1.22	0.156	−0.129		1.247
0.26	1.24	0.165	−0.175		1.23
0.28	1.26	0.173	−0.214		1.219
0.30	1.27	0.179	−0.245	0	1.204
0.32	1.28	0.183	−0.271	−0.049	1.143
0.34	1.29	0.188	−0.293	−0.090	1.095
0.36	1.30	0.191	−0.311	−0.123	1.057
0.38	1.31	0.194	−0.325	−0.150	1.029
0.40	1.32	0.196	−0.336	−0.173	1.01

DESIGN

The foregoing analysis is readily extended to design. There are two basic design problems in the theory of feedback control systems: the adjustment of gain and the determination of appropriate values for circuit parameters. Both facets of design are conveniently illustrated in terms of the simple example used in the preceding section, with

$$G(s)=\frac{10K}{s(s+10)} \quad \text{and} \quad T=0.1 \text{ second} \tag{6}$$

A trial-and-error procedure suffices for the rapid adjustment of the gain K to meet any of the customary specifications on relative stability, rise time, etc. For example, if the desired overshoot is specified, a value of K is chosen and the $c(t)$ curve is developed until the first maximum is reached. If the overshoot is too large, a smaller value of K is chosen and the analysis repeated. With a little experience the proper value of K can be established with one or two attempts. With K set at 1.34 the example of the preceding section demonstrates the solution with the specified overshoot about 29 per cent.

Insight into the more general problem of fixing system parameters can be gained from a study of the effect of variation of aT on the response of a system with a sampling period T and

$$G(s)=\frac{aK}{s(s+a)} \tag{7}$$

If a maximum value of overshoot is specified (25 to 30 per cent in this paper), an appropriate value of K and the corresponding form of the response can be determined for any value of aT.

If aT has a value of 5, the $g(t)$ response has essentially reached steady state before the next sampling instant. Parts A, B, and C of Fig. 5 show the nature of $c(t)$

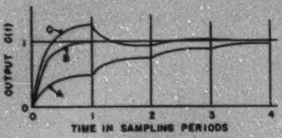


Fig. 5. Response of system with aT=5 as K is varied

Curve A—K=0.5
Curve B—K=1
Curve C—K=1.25

as K is varied. For most purposes the system of part B with K equal to unity would be preferred. The only advantage of the system of part C is that it has a slightly shorter rise time, but this is certainly outweighed by the oscillatory nature of the response. For this reason the value of gain is ordinarily never set higher than the magnitude to put the $c(t)$ curve through unity at the end of one sampling period.

The results of continuing this investigation for aT equal to 2.5, 1.5, 1.0, and 0.5 are shown in parts A, B, C, and D respectively of Fig. 6. As aT is decreased the magnitude of $c(t)$ after one period decreases and more sampling periods are required for the $c(t)$ curve to reach its maximum value. This is shown in Table II which also includes the value of $c(T)$ and K when K assumes the value placing the system on the verge of instability. The equations for $K_{\max}$ and $c(T)_{\max}$ are derived in Appendix II and are

$$K_{\max}=\frac{2(1+\epsilon^{-aT})}{1-\epsilon^{-aT}} \tag{8}$$

$$c(T)_{\max}=2(1+\epsilon^{-aT}) \tag{9}$$

SYNTHESIS

When more freedom is allowed the designer and more comprehensive problems are considered, it becomes necessary to shape the response characteristics to meet arbitrary specifications. In terms of the design of sampled-data systems, synthesis involves the reshaping of the impulse response $g(t)$.

This reshaping of a time function can be accomplished in a variety of ways ranging from almost pure synthesis to simply extended analysis. Idealistically, the designer simply forces the system to behave in the desired manner. The unwanted poles and zeros of $G_2(s)$, the transfer function of the controller system, are canceled by zeros and poles of $G_1(s)$, and the critical frequencies reinserted are selected to yield a system which meets the performance specifications.

If the system required to effect such a straightforward compensation is undesirably complex, the impulse response can be shaped more indirectly. For example, $G_2(s)$ conventionally includes a pole at the origin: i.e., $G_2(s)$ can be written as $T(s)/s$. The impulse response $g_2(t)$ is then simply the step-function response of a system with the transfer function $T(s)$. Hence, all the literature and knowledge concerning the effects of various pole-zero configurations can be brought to bear on the problem of reshaping $g_2(t)$ or the step-function response for $T(s)$.

If the designer leans even more heavily away from synthesis and toward design by analysis, the poles of $G_2(s)$ can be modified one by one until the appropriate response is obtained. For example, a typical design problem might be to decrease the rise time and settling time while leaving the overshoot unchanged. The approach is illustrated with the specifications

$$G_2(s)=500K/s(s+10)(s+50) \text{ with } T=0.05 \text{ second} \tag{10}$$

Since it is the pole at -10 which is essentially limiting the rise time, it is clear that $G_1(s)$ should effectively move this pole farther to the left in the s-plane. If $G_1(s)$ is

$$G_1(s)=\frac{5(s+10)}{s+50} \tag{11}$$

the pole is moved to -50, and $G(s)$ becomes

$$G(s)=\frac{2{,}500K}{s(s+50)^2} \tag{12}$$

With K set to meet the overshoot requirements the curve labeled part B of Fig. 7 results, in contrast to the curve of part A for the response without compensation. More stringent compensation specifications can be met by moving the pole at -10 farther to the left with a

$$G_1'(s)=\frac{10(s+10)}{s+100} \tag{13}$$

Part C of Fig. 7 shows the result after K is adjusted to make $c(T)$ equal to unity, in which case the overshoot is 17 per cent.

With a little experience the designer can examine $G_2(s)$ and T and in most cases see the nature of the needed $G_1(s)$. If the $g_2(t)$ is monotonic increasing, specific statements regarding the desired compensation can be made. If $g_2(t)$ essentially reaches steady state in one sampling period, little needs to be done in the way of compensation. If this is not true, the pole nearest the origin is predominant in determining the rise time and is the one to move first. If there are several poles near the origin, the movement of any one of these to the left decreases rise time. If the resulting system does not satisfy the specifications, additional poles must be moved. Although relations similar to equations 8 and 9 cannot be derived for the general monotonic increasing case, it can be stated that $K_{\max}$ and $c(T)_{\max}$ are both equal to or greater than 2. If a system is designed with $c(T)$ equal to or less than unity, the system is operating well below the limit of stability and each successive maximum value of $c(t)$ is lower than the preceding.

If $g_2(t)$ is oscillatory, general statements are more difficult. In certain cases the gain for absolute stability is lowered as a result of the oscillation, but in other cases, even though the gain for absolute stability is not affected, K has to be set at a lower value to meet the overshoot requirement and the entire response is more sluggish.

The curve of the desired $g(t)$ is one

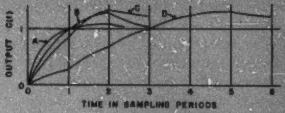


Fig. 6. Response of system as aT is varied

Curve A—aT=2.5, K=1.09
Curve B—aT=1.5, K=1.29
Curve C—aT=1.0, K=1.34
Curve D—aT=0.5, K=0.77

Table II. Design Summary

aT	Per Cent Overshoot	K	$K_{\max}$	c(T)	$c(T)_{\max}$	Periods to Maximum
5	0	1	2	1	2.03	1
2.5	8	1.09	2.36	1	2.16	2
1.5	22.5	1.29	3.15	1	2.45	2
1.0	29	1.34	4.32	0.846	2.74	2
0.5	26.9	0.77	8.18	0.303	3.22	5

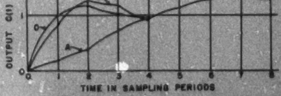


Fig. 7. Response of system being synthesized

which reaches the range of 80 to 100 per cent of its steady-state value in one sampling period and settles down to the steady state rapidly. In most cases an examination of $g_2(t)$ indicates the nature of the required compensation network.

Limitations of the Impulse Response Approach

When the z-transformation theory is used to determine stability the function $G(z)/[1+G(z)]$, where $G(z)$ is the z-transform of $G(s)$, is examined and the system becomes absolutely unstable when a pole or poles move out of the unit circle in the z-plane. If the first pole leaves the unit circle on the negative real axis (corresponding to system oscillation at one-half the sampling frequency, or oscillation basically because of the sampling), the impulse response approach can be used to determine the maximum gain. However, if the first pair of poles leave the unit circle at some point other than on the negative real axis, it is impossible to determine the maximum gain directly except in special cases. (This mode of instability can happen in a second-order system with a dead-time delay or a holding circuit). The only method left to the designer using the impulse response approach is to set the gain and follow the output long enough to determine system stability.

A similar situation occurs in attempting to find the steady-state output when the input to the system is a sinusoidal function. When the z-transformation is used the envelope of the output and the bandwidth can be found directly. If the impulse response approach is used and the sampling period is not a subharmonic of the sinusoidal period, the $c(t)$ curve never repeats, and it is difficult to recognize steady-state conditions. If the two methods of analysis are combined, the envelope can be found by z-transformation methods and the output between sampling instants by the impulse response approach, giving the complete picture of the output.

Conclusions

The present z-transformation theory suffers from the uncertainty surrounding the behavior of the system between sampling instants. This is partly removed by the procedure of finding the output at submultiples of the sampling period and is entirely removed by the impulse response approach. The z-transformation method gives a comprehensive picture of absolute stability, but this is not the problem faced by the engineer who is trying to meet such specifications as maximum overshoot. The impulse response approach can be used to set the gain to meet overshoot specification and an examination of $g_2(t)$ in most cases indicates the nature of the compensation network needed to meet the specifications. Thus the optimum design method for a large variety of sampled-data systems is a design on the basis of a shaping of the impulse response of the open-loop system. The z-transformation and Nyquist diagram are used to increase the designer's circumspection.

Table III. z-Transform

Time Function	Laplace Transform	z-Transform
$u_0(t)$	1	1
$u_{-1}(t)$	$\frac{1}{s}$	$\frac{z}{z-1}$
t	$\frac{1}{s^2}$	$\frac{Tz}{(z-1)^2}$
$\epsilon^{-\alpha T}$	$\frac{1}{s+\alpha}$	$\frac{z}{z-\epsilon^{-\alpha T}}$

Appendix I. Output Between Sampling Instants

The sampled functions are described throughout this paper in terms of the corresponding z-transforms. The definition of the z-transform follows that used by Ragazzini and Zadeh.[1] The z-transform is obtained by substituting z for ϵ^{sT} in the Laplace transform of the sequence of samples. As a consequence of this definition, the expansion of $C(z)$, in a power series in $1/z$, places in evidence the successive sample amplitudes of $c(t)$. $C(z)$ is obtained in closed form with the aid of Table III. (The Ragazzini-Zadeh article contains a more extensive list of transforms.) Conversion from $C(s)$ to $C(z)$ is accomplished by a partial fraction expansion of $C(s)$, followed by a term-by-term identification based on the table.

Thus the z-transformation can be used to evaluate rapidly the output of a sampled-data system at the sampling instants. In this basic application the z-transform $C(z)$ contains no information about the behavior of the actual $c(t)$ between sampling instants. The artifice of the introduction of a fictitious sampler permits the circumvention of this disadvantage and allows evaluation of the output at submultiples of the sampling period.

In a closed-loop system with the sampler following the error-measuring device, the actuating signal $e^*(t)$ is a train of impulses at the sampling frequency and is zero between sampling instants. The z-transform of the actuating signal can be expanded in the form

$$E(z)=\frac{R(z)}{1+G(z)}=A_0+\frac{A_1}{z}+\frac{A_2}{z^2}+\frac{A_3}{z^3}+\ldots \quad (14)$$

The time between successive values is T.

As shown in Fig. 8, a fictitious sampler is added in tandem with the actual sampler. The period of the fictitious sampler is assumed a submultiple of the period T.

$$T'=\frac{T}{n} \quad (15)$$

In this example, if n is given a value of 2, the actuating signal at the output of the fictitious sampler is represented by

$$E'(z)=A_0+\frac{0}{z}+\frac{A_1}{z^2}+\frac{0}{z^3}+\frac{A_2}{z^4}+\frac{0}{z^5}+\frac{A_3}{z^6}+\ldots \quad (16)$$

where the time between successive values corresponds to T' and the A's are the same as in equation 14. $E'(z)$ can be written as

$$E'(z)=\frac{R(z^2)}{1+G(z^2)} \quad (17)$$

and the z-transform of the output $C'(z)$ as

$$C'(z)=\frac{R(z^2)}{1+G(z^2)}G'(z) \quad (18)$$

in which $G'(z)$ is the z-transform of $G(s)$ with respect to T' and $R(z^2)$ and $G(z^2)$ are the z-transforms of $R(s)$ and $G(s)$ with respect to T, but with z replaced by z^2.

$G(s)$ from equation 3 serves as an example

$$G(s)=\frac{13.4}{s(s+10)}=1.34\left(\frac{1}{s}-\frac{1}{s+10}\right) \quad (19)$$

The corresponding $G(z)$ is given from Table III

$$G(z)=1.34\left(\frac{z}{z-1}-\frac{z}{z-\epsilon^{-1}}\right)=\frac{0.846z}{z^2-1.368z+0.368} \quad (20)$$

With the input a unit step function, $E(z)$ is

$$\frac{R(z)}{1+G(z)}=\frac{z^2-0.368z}{z^2-0.522z+0.368} \quad (21)$$

and

$$E'(z)=E(z^2)=\frac{z^4-0.368z^2}{z^4-0.522z^2+0.368} \quad (22)$$

$G'(z)$ is given by

$$G'(z)=1.34\left(\frac{z}{z-1}-\frac{z}{z-\epsilon^{-0.5}}\right)=\frac{0.526z}{z^2-1.607z+0.607} \quad (23)$$

The output is described by the z-transform

$$C'(z)=\frac{0.526z^5-0.194z^3}{z^6-1.6072z^5+0.085z^4+0.838z^3+0.052z^2-0.592z+0.223} \quad (24)$$

When equation 24 is expanded by

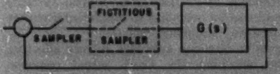


Fig. 8. Simple system with fictitious sampler

ordinary long division of the denominator into the numerator, the first few terms are

$$C'(z)=0+\frac{0.526}{z}+\frac{0.846}{z^2}+\frac{1.12}{z^3}+\frac{1.29}{z^4}+\ldots \tag{25}$$

This result agrees with the curve of $c(t)$ in Fig. 4.

This procedure for determining the response between sampling instants is straightforward. The only difficulty arises in the evaluation of $G'(z)$ from the known $G(s)$ or $G(z)$. $p_{G'}$, a pole of $G'(z)$, is related to the corresponding pole p_G of $G(z)$ by the equation

$$p_{G'}=(p_G)^{1/n} \tag{26}$$

The residues of $G(z)$ and $G'(z)$ are equal in the corresponding poles. Unfortunately, however, there is no simple correlation between the numerator polynomials of $G(z)$ and $G'(z)$. Except in simple cases, the numerator of $G'(z)$ must be evaluated by summation of the terms in the partial fraction expansion.

Appendix II. Stability of Second-Order System

Equations 8 and 9 are derived in the following manner

$$G(s)=\frac{aK}{s(s+a)}=\frac{K}{s}-\frac{K}{s+a} \tag{27}$$

The z-transformation of equation 27 is

$$G(z)=\frac{Kz}{z-1}-\frac{Kz}{z-\epsilon^{-aT}}=\frac{K(1-\epsilon^{-aT})}{z^2-(1+\epsilon^{-aT})z+\epsilon^{-aT}} \tag{28}$$

which gives

$$\frac{G(z)}{1+G(z)}=\frac{K(1-\epsilon^{-aT})z}{z^2+K(1-\epsilon^{-aT})-(1+\epsilon^{-aT})z+\epsilon^{-aT}} \tag{29}$$

As one step of the z-transformation the change of variable

$$z=\epsilon^{sT} \tag{30}$$

is made to map the left half of the s-plane to the interior of the unit circle in the z-plane. As K is increased the system becomes unstable when the poles of $G(z)/[1+G(z)]$ leave the unit circle. Thus stability requires that the zeros of

$$p(z)=z^2+K(1-\epsilon^{-aT})-(1+\epsilon^{-aT})z+\epsilon^{-aT} \tag{31}$$

lie within the unit circle. The system is stable if

$$p(0)<1 \tag{32}$$

$$p(1)>0 \tag{33}$$

$$p(-1)>0 \tag{34}$$

Equation 32 gives

$$p(0)=\epsilon^{-aT}<1 \tag{35}$$

which is true for all aT greater than zero and for all K. Equation 33 yields

$$p(1)=K(1-\epsilon^{-aT})>0 \tag{36}$$

which is true for all positive aT and for all K. Finally, equation 34 gives

$$p(-1)=-K(1-\epsilon^{-aT})+2(1+\epsilon^{-aT})>0 \tag{37}$$

from which K_{max} or equation 8 can be derived.

The inverse Laplace transform of equation 27 yields

$$g(t)=K(1-\epsilon^{-aT}) \tag{38}$$

When K_{max} is substituted and $g(T)$ evaluated, equation 9 results.

Equations 32, 33, and 34 are valid for the second-order system only. In the general case, specific procedures are available for establishing stability in terms of the z-transform.[2]

References

1. THE ANALYSIS OF SAMPLED-DATA SYSTEMS, J. R. Ragazzini, L. A. Zadeh. *AIEE Transactions*, vol. 71, pt. II, Nov. 1952, pp. 225–34.

2. THE GEOMETRY OF THE ZEROS OF A POLYNOMIAL IN A COMPLEX VARIABLE (book), M. Marden. American Mathematical Society, New York, N. Y., 1949, p. 152.

◆

Discussion

Eliahu I. Jury (University of California, Berkeley, Calif.): This paper introduces a different design technique (in contrast to the frequency approach) for synthesis of sampled-data control systems. The impulsive response method of design presented is at best applicable to special and simple cases of sampled-data systems. For instance, in cases where components exist in the feed-back path or where different types of sampled-data control systems[1] are being investigated, or where the uncompensated system is of higher order than the second, it appears that the method presented is not conveniently or readily applicable. In such cases, the frequency response method seems to be more promising since the design information obtained for second-order systems can be generally extended to higher order and more complicated systems.

Although the critical study of the system response between sampling instances is desirable, it is not very essential in the design of sampled-data systems, for in most cases the designer is primarily concerned with the response at the sampling instances. The z-transform approach yields the transient as well as the steady-state response of sampled-data control systems and can be extended to obtain information between sampling instants.[2,3] The impulsive response approach is very effective in obtaining the exact actual response between sampling instances when such a critical study is required.

As indicated in the paper, the impulsive response approach is not readily applicable for determining the maximum gain when the poles of $G^*(z)/1+G^*(z)$ moves out of the unit circle at some point other than on the negative real axis. However, the root locus plot of the characteristic equation $1+G^*(z)$ readily yields the maximum gain for the chosen parameters. Furthermore, the root-locus plot in the z-domain gives an indication of how to compensate such systems to obtain better performance. Fig. 9 illustrates the application of the root-locus for a chosen system. This system is similar to the example given in Appendix II with the exception that a zero-order hold circuit is utilized.

It is of importance to mention that to minimize the output ripple which is produced by the sampling process, it is desirable to use a hold circuit or low-pass filter in the system loop. This requirement often necessitates the first pair of the complex poles of the system function moving out of the unit circle in the z-plane other than on the negative real axis.

In conclusion, I feel that the z-transform method is very useful in obtaining a deep insight into the behavior of sampled-data control systems and very effective in the design and synthesis of such systems. The z-transform approach combines the transient and the steady-state performance and finally can be extended to obtain the required information between sampling instances. The impulse response approach is very informative of the behavior of sampled-data systems for simple cases, and the contribution of this approach for complementing the z-transform method is indeed very valuable.

REFERENCES

1. See Table II in reference 1 of the paper.

2. AUTOMATIC AND MANUAL CONTROL (book)

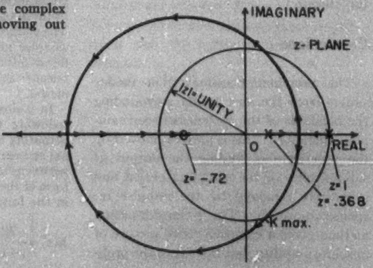


Fig. 9. Root locus of $1+G^*(z)$ for

$$G(s)=\frac{K(1-\epsilon^{-Ts})}{s^2(s+a)},\ a=1,\ T=1$$

$$G^*(z)=\frac{0.368K}{z-1}\left[\frac{z+0.72}{z-0.368}\right]$$

Butterworths Scientific Publications, London, England, 1952. Discussion by R. H. Barker, pp. 404–07.

3. EXTENSION OF CONVENTIONAL TECHNIQUES TO THE DESIGN OF SAMPLED-DATA SYSTEMS. W. K. Linvill, R. W. Sittler. *Convention Record of I.R.E.*, Institute of Radio Engineers, New York, N. Y., pt. I, 1953, pp. 99–104.

John M. Salzer (Santa Monica, Calif.): This paper deserves commendation for its instructional value, as well as for the simplicity and lucidity of its approach. It may be criticized for some of its sweeping statements and for the meagerness of its references. The major contribution of the paper is the heuristic insight it gives regarding the nature of certain sampled-data systems.

The authors select a simple system configuration, which has become all but canonical for papers dealing with sampled-data systems. They introduce the reader to the understanding of the behavior of such a system, which should prove very illuminating to the beginner. The straightforward approach is then utilized to yield criteria for the adjustment of gain and the nature of forward-loop compensation. One must not be misled by the simplicity of the approach because the method is very useful where the design is to be done in terms of the transient behavior of such a system to a specific input. The examples given by the authors aid importantly in clarification.

More reference to actual systems would have stressed the practicality of the analysis. One notes that the output of the sampling switch is directly applied to the plant, $G_2(s)$, except where compensation is inserted. Thus, the operation is at a very low duty cycle and subject to saturation effects. The often-used method of clamping the output of the sampling switch is not used in the examples. There is no objection to the absence of the clamping unit where a compensating network, $G_1(s)$, is utilized, since it acts as a sort of holding device. In view of the lack of reference to particular applications, it is not clear what is meant by such statements as: "...the value of gain is ordinarily never set higher than the magnitude to put the $c(t)$ curve through unity at the end of one sampling period."

In the conclusion the authors note that Nyquist diagrams can yield additional circumspection in the design procedure. Some references to papers illustrating this method in addition to reference 1 of the paper would have been in order. Even more helpful would have been a demonstration at least in the examples of what additional circumspection the Nyquist diagrams could yield. The use of the "fictitious" sampling at twice (or n-times) the rate of the actual sampling device illustrated in Appendix I could have been correlated with previous work[1] along these lines. The use of the so-called z-transform method is well illustrated, and it is clear that the impulse response, $g(t)$, rounds out the general picture. Nevertheless, the method of design illustrated is essentially the z-transform method. It is particularly noted that in all examples $g(t)$ is a monotonic function so that all information concerning overshoot and oscillatory behavior of the system is completely contained in the samples themselves, and the need for the impulse response is questionable. On the other hand, when the impulse response is oscillatory, general statements are more difficult, and in this case the utility of the impulse response method may become questionable.

REFERENCE

1. See reference 3 of Mr. Jury's discussion.

Study of Transformerless Rectified Higher Voltage D-C Aircraft Electric System

J. P. DALLAS
MEMBER AIEE

C. A. REISING, JR.
MEMBER AIEE

Synopsis: Variable-frequency a-c systems and paralleled d-c supplies operating simultaneously from the same a-c generators and with the elimination of transformers, voltage regulators, and reverse-current relays from the d-c supplies are proposed as a means of increasing aircraft reliability and of reducing costs, complexity, and weight. The suggested system can be connected to provide either a single direct voltage or two direct voltages obtained without transformers and without interfering with the simultaneous use of the a-c generators for conventional variable-frequency a-c loads.

SIMPLICITY is the one common denominator for reducing aircraft cost, unreliability, weight, flight casualties, and nonoperative airplanes in a flight group. Yet aircraft electric systems in the United States are increasing in complexity. This would seem to justify a careful re-examination of many basic concepts to determine if aircraft electric system simplicity is entirely a lost cause.

Simplicity is not a function of aircraft size. A large airplane with a d-c single-conductor main electric system may be less complex than a smaller airplane with a 400-cycle 3-phase 3-wire grounded-neutral constant-frequency main power generating system requiring complicated hydromechanical constant-speed generator drives. Similarly, a calendar watch may be more complicated than a locomotive but, if simplicity is an objective, perhaps separating the functions of hourly and calendar time would improve the reliability and lower the cost of the watch. A somewhat analogous separation of aircraft electric power systems on the basis of function will be discussed in this paper.

Simplicity is not excluded by the airplane-type mission or load requirements. Expanding concepts of the purpose or mission of aircraft and the addition of new types of electric loads make changes inevitable. In the interest of simplicity, major a-c loads require a-c generating systems; and on very large airplanes voltages higher than 28 volts are unavoidable. However, this does not mean that the present complexity of the basic aircraft electric system is unavoidable.

Simplicity rigorously and unremittingly pursued may still be a practical objective. We suggest that the prospects of its achievement may be bettered by the following:

1. A continuous drive effort to reduce the number of parts, the number of circuits, the complexity of each circuit, and the number of interdependent circuits or devices where a failure in one circuit will affect the operation of several. The drive for optimum simplicity must grow from the initial studies of the system to the completed installations, from the first block diagram to the finished airplane.

2. The courage at least to consider every avenue of simplification including those which may require some sacrifice of standardization, and/or the minor modification of load equipment, if such changes would result in significant over-all better performance of the airplane in achieving its mission. Everyone is in favor of simplification but, when it comes to even the smallest sacrifice of standardization or the modification of load equipment, a cautious attitude is common.

Paper 54-309, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted June 29, 1953; made available for printing April 28, 1954.

J. P. DALLAS and C. A. REISING, JR. are with the Aircraft Division, Hughes Tool Company, Culver City, Calif.

3. Electric system performance limits or tolerances should be considered on the basis of over-all aircraft simplicity and performance as opposed to the acceptance without question of currently published limits. This requires rigorous analysis of the *real limits* defining acceptable aircraft system performance for the proper functioning of load equipment. The word *real* is emphasized because often expensive and complicated close-tolerance limits on frequency, harmonic content, voltage, etc., are specified, because the *necessary investigation has not been made to establish the real limits, and prior to such investigation it is necessary to stay on the "safe side."* The Naval Research Laboratory illustrates this point in the recent investigation of power-wave form requirements necessary for satisfactory operation of present radar loads in which it was found that "standard airborne radar can operate satisfactorily on complex wave power inputs which greatly exceed the tolerances presently allowed by existing military specifications."[1]

In general, the principal cause of nonessential complexity in aircraft systems has come from the not always essential close-tolerance requirements on voltage, frequency, harmonic content, etc. Often a close-tolerance limit that is a small advantage to load equipment means critical increase in the complexity of the main power-generating system. The principle of isolating and reducing to the minimum all loads requiring close-tolerance power may allow such power to be economically supplied by ancillary sources and thereby to result in lower dollar and weight cost for the airplane by simplifying the main bulk power source.

A simplification chart to accompany the electric load chart in preliminary aircraft design is suggested as an aid to design personnel for focusing attention continuously on this electrical simplification problem. One useful form of this chart might show an evaluation of weight and a reliability and simplicity factor for each electrical item on the airplane for each system under consideration. This factor may be estimated in any number of ways. A figure based on the total number of circuits and component parts for the item would be a rough and ready method to use. A correction factor of a 25-per-cent decrease in reliability could logically be added to any item using vacuum tubes, based on current reliability experience with this type of equipment. Any number of variations on this idea could be useful, and the suggestion is made in the belief that any systematic and conscientious effort to evaluate the simplification problem in initial aircraft electrical design can prove of value.

Finally the achievement of optimum simplicity in aircraft electric systems is the business of arriving at the most judicious compromise between requirements for load equipment, main electric system simplicity, and standardization for the benefit of over-all aircraft performance. Pursuing this formula for the simplification of an aircraft electric system the following study of a transformerless rectified higher voltage d-c aircraft electric system resulted.

Combined Variable-Frequency A-C and Higher Voltage D-C Generating System

It may be practical to join together a standard 320- to 1,000-cycle 115/200-volt 3-phase variable-frequency aircraft a-c generating system with a transformerless half-wave 3-phase rectifier d-c supply, so that both a-c and d-c loads can be supplied from the same a-c generator. By connecting the d-c output of several such generator-rectifier a-c/d-c supplies in parallel, the reliability of a parallel-source d-c system is obtained. Parallel-operated power sources from propulsion power plant drives are deemed essential for the minimum reliability requirements of many present aircraft. Since present aircraft designs are requiring increasing amounts of a-c power beyond that economically obtainable from inverters, the simplest solution seems to be to generate it. The possibility of obtaining a-c bulk power in its simplest form, and also parallel-operated sources of high-reliability d-c power from standard generators without transformers or complicated direct-voltage regulator systems, should be of enough interest to merit substantial study. Of course there are problems, and perhaps insurmountable difficulties, in the way of adapting such a system to aircraft needs. Only study, discussion, and tests will tell.

OUTLINE OF SYSTEM

Referring to the block diagram in Fig. 1, the following is noted:

1. The a-c generators are driven from the airplane's propulsion power plant and are standard 115/200-volt 320-to-1,000-cycle variable-frequency units wye-connected with the airplane structure acting as the grounded neutral. Since the frequency of the a-c system is allowed to vary with the speed of the propulsion power plant, no complicated constant speed drive is needed. The complete a-c generating system including exciter, voltage regulator, and generator circuit protection system is standard in every respect, as are the a-c load protection and changeover provisions. Each a-c generator has an additional load; the d-c loads supplied through the selenium rectifier.

2. The selenium rectifier is connected directly, without transformers, to the a-c generator load bus. The usual load circuit protection provisions are not shown, but of course are not omitted. The d-c output of the rectifier feeds the d-c bus, operating in parallel with one or more similar units. In Fig. 2, a simplified diagram of the rectifier-generator d-c bus connection is shown. A 3-phase half-wave circuit is used. It is notable that no provision for separate direct-voltage regulation is called for.

3. Both transformer and direct-voltage regulator are eliminated. The d-c circuit is simplified by elimination of both the usual transformer and direct-voltage regulator. Omission of the transformer between the a-c generator and the rectifier may make possible the elimination of the direct-voltage regulator by eliminating the transformer voltage regulation factor. The voltage drop in selenium rectifiers is very low. For the new 40-rms volt cells it is only 1.11 volts per cell at full load and for 33-volt units 0.98 volt. This would give a total full-load voltage drop in the rectifier of as low as 6.66 volts for the 40-volt cells and approximately 6.86 volts for the 33-volt cells.

The approximately 6-volt rectifier-voltage regulation must be added, of course, to the specified voltage regulation of the a-c generator system, since there is no direct-voltage regulator. The a-c regulation, as specified in *MIL-E-7894*,[2] is 18 volts which would bring the total d-c regulation to approximately 24 volts for any steady-state condition between no load and full load. This may exceed desirable d-c system limits but it is not beyond the possibility of practical operation. Voltage regulation will be discussed in the section entitled "Disadvantages of System." However, the regulation to be expected of aircraft transformers is not as good as that of the selenium rectifier. Five to 7 per cent transformer regulation is standard, and this would have to be added to the regulation factors noted previously, unless a direct-voltage regulation system is used. The elimination of the usual transformer, also, can be counted on to increase the efficiency of the d-c supply by at least 10 per cent. The additional saving of approximately 5 per cent, by eliminating the reactors used for direct-voltage regulation, makes probable a total gain in efficiency of approximately 15 per cent for the d-c transformerless supply.

4. Eliminating the transformer fixes the nominal direct-voltage at 126 for a 3-phase half-wave rectifier connected to the standard 115/200-volt wye-connected aircraft a-c system. With the new 40-rms volt selenium cells only 4 cells per phase would be required. The schematic is shown in Fig. 2.

5. Voltage regulation and load sharing is suggested. It is proposed that no voltage-regulation equipment be used in the d-c system. The low-voltage drop of the selenium rectifier of 6 to 7 volts at full load may make this possible. The voltage regulation for the a-c system is intended to be the standard voltage-regulation system normally used with the particular variable-frequency a-c generators employed. The alternating-voltage regulators would have to be adjusted for sufficient voltage droop

with load, to ensure a practical degree of load sharing of the d-c load among the a-c generators.

The voltage limits of 102- to 124-rms volts line to neutral specified in *MIL-E-7894*[2] should be sufficient to allow a practical degree of load balance. Load balance in the d-c supplies would be obtained as the result of the combined a-c generator and rectifier voltage with load droop. It is to be noted, of course, that no load sharing of the a-c loads which are not from paralleled sources is possible or intended. The a-c loads are in every case associated with a particular a-c generator and, if stand-by provision for such loads are required then load transfer provision must be made. Only the loads reflected into the a-c system via the paralleled d-c rectifier outputs are considered in the foregoing. To ensure some practical minimum degree of load balance among the a-c generators, strictly for these rectifier loads, a substantial alternating-voltage droop with load of the a-c generators was suggested. While the resulting load balance among the various machines may be less than the optimum desired, the added weight in extra generator capacity, to compensate for less than optimum load balance, may be less than the weight of the additional voltage regulation and load equalizing provisions of a conventional system, and the gain in simplicity, reliability, and cost reduction may be substantial.

To summarize this proposal, regard the a-c portions in Figs. 1 and 2, including all alternating-voltage regulation, circuit protection, and load functions, as standard in every respect. The proposal is to superimpose on this standard variable-frequency a-c system a 3-phase half-wave d-c selenium-rectifier supply on each a-c generator, and to connect the outputs of the several such d-c rectifier supplies in parallel to a common d-c bus.

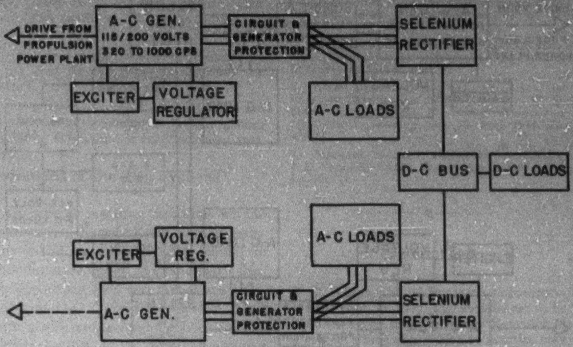

DRIVE FROM PROPULSION POWER PLANT
A-C GEN. 115/200 VOLTS 320 TO 1000 CPS
CIRCUIT & GENERATOR PROTECTION
SELENIUM RECTIFIER
EXCITER
VOLTAGE REGULATOR
A-C LOADS
D-C BUS
D-C LOADS
A-C LOADS
EXCITER
VOLTAGE REG.
A-C GEN.
CIRCUIT & GENERATOR PROTECTION
SELENIUM RECTIFIER

Fig. 1. Block diagram of transformerless a-c/d-c aircraft electric system

Disadvantages of System

Wave form distortion of the a-c supplies caused by the half-wave 3-phase rectifier loads may prove to be the major disadvantage of this proposal. Obviously, the a-c system must be capable of supplying all of the loads normally supplied by variable-frequency a-c generators. Electronic and heating loads will be the major items. Preliminary experimental tests on a small model system of 1-kva capacity which showed a 13-per-cent third harmonic at full-rated a-c load developed 30-per-cent second harmonic content and a 2-per-cent third harmonic with full-rated d-c 3-phase half-wave rectifier loading. The 13-per-cent third harmonic for the a-c load was, of course, far in excess of specified limits, but is actually not untypical of many aircraft a-c power supplies including many inverters. The increased wave distortion with the d-c rectifier loading was a function of the ratio of d-c rectifier to a-c loading of the machine. With the machine full-load output split equally between the rectifier and an a-c resistive load the harmonic content dropped to 18-per-cent second harmonic and 7-per-cent third harmonic.

The Naval Research Laboratory has recently investigated the wave form requirements necessary for satisfactory operation of some aircraft electronic loads.[1] It was found that "standard airborne radar can operate satisfactorily on complex wave power inputs greatly exceeding the tolerances presently allowed by existing military specification." It is not known whether much present-day electronic equipment could tolerate wave form distortion of the magnitude aforementioned, but it is suggested that, if such distortion should prove an obstacle to achieving a simplified aircraft electric system, perhaps isolating those functions requiring better wave form and supplying internal filters should be considered.

Eighteen-per-cent ripple in a single half-wave 3-phase rectifier output is to be expected. Where several rectifiers are feeding a common d-c bus, since neither the phase relationship nor the frequency is relatively controlled, this ripple factor should be much less than 18 per cent. It may be recalled that some of the commercial rectifier cart ground power supplies now employed on present 28-volt d-c airplanes use 3-phase half-wave rectification and have 18-per-cent ripple. The authors cannot recall any malfunction of present d-c airplanes from this 18-per-cent ripple. A careful check of the electrical item list of several typical airplanes uncovered only one item which was regarded as questionable in this connection, i.e., sensitive polarized relay circuits. An operation check of such a sensitive relay device using a Barber-Coleman relay adjusted to a sensitivity of one milliwatt showed, surprisingly, a slight improvement of operation as a result of operating the whole unit on a d-c rectifier system with an 18-per-cent ripple. The slight variation in voltage was found to produce more accurate trip points of the sensitive elements.

D-c system voltage limits will range from a low of 112 volts to 142 volts absolute maximum for a-c system voltage limits of 102 to 124 volts specified in *MIL-E-7894*.[2] This is a + 13-per-cent

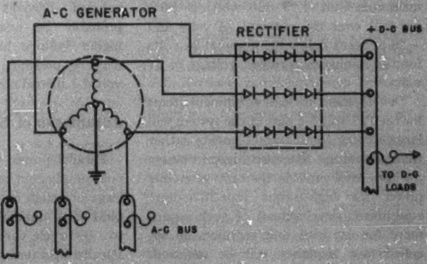


Fig. 2. Simplified schematic diagram of a-c/d-c transformerless system, circuit protection omitted

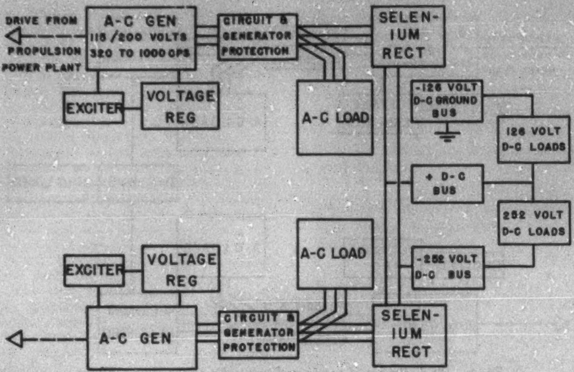


Fig. 3. Block diagram of a transformerless a-c/d-c system supplying two direct voltages

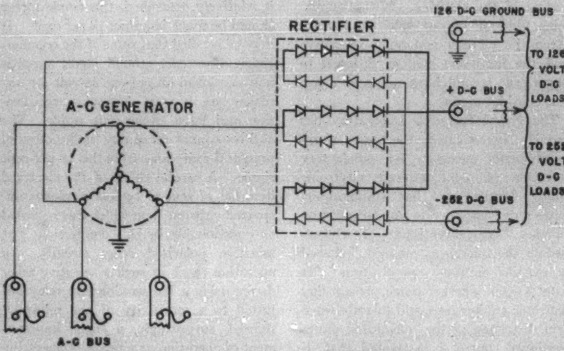


Fig. 4. Simplified schematic diagram of a-c/d-c transformerless system supplying two direct voltages, circuit protection omitted

and −11-per-cent voltage-limit range from the 126 nominal voltage. Since there are no currently active higher voltage d-c specifications; the present *MIL-C-7894* voltage limits on the nominal 24 volt system is compared as to percentages. This system is specified for a maximum limit of 29 volts which is +21-per-cent over the nominal 24 volts and −29-per-cent below the nominal 24 volts for the minimum specified voltage of 17 volts.

Development of load equipment items for the 126-volt d-c part of the system will largely come about from necessity rather than planning. Because higher voltage d-c may well provide the only economic power for high-torque repetitive-start equipment, development of such equipment for use with unit rectifiers on so-called a-c airplanes will be required. Cargo handling and hoisting, intermittent actuator operation, ammunition charging and boosting, and other high-torque repetitive-start motor applications represent a major portion of the total electric loads, exclusive of possible deicing applications on many airplanes. Presently developed 120-volt d-c equipment can probably be adapted to the slightly higher voltage here considered and it may be possible to import much 125-volt d-c British equipment.

Advantages of System

Parallel power source reliability without complicated constant-speed drives or large inverters is obtained from the simplest of all standard a-c systems: the 320- to 1,000-cycle variable-frequency system. Parallel-operated power sources are considered essential for many loads on present airplanes, especially where propulsion power plants are used for driving the generators.

Elimination of the transformer, voltage regulator, and reverse-current relay from conventional rectified d-c systems is expected to greatly simplify, reduce weight, and substantially lower the cost of the system. While some problems arise, as outlined in "Disadvantages," a net gain for the over-all operation of the airplane may result.

D-c parallel-power source without brushes or moving parts is of course an advantage common to all rectifier supply systems but is secured perhaps more simply in this proposal.

An ideal electric power for cargo handling and other high-torque or repetitive-start loads is provided by the 126-volt d-c supply.

A-c and d-c power supplies from the same variable-frequency generator with only one standard voltage regulator is used for the combined systems.

Improved efficiency of the d-c system as compared to conventional rectifier systems should result from the elimination of transformer and direct-voltage regulator functions. Efficiency of the complete d-c conversion system should be better than 85 per cent at full load and at least 80 per cent for half load. Conventional rectifier systems for converting a-c generator supplies to d-c, using transformers and direct-voltage regulation controls, are likely to be from 65-per-cent to 70-per-cent efficient at full load and may drop to as low as 50 per cent at half load.

Combined Variable-Frequency A-C and Full-Wave 126/252-Volt D-C Generating System

The same basic transformerless system described previously can be carried one step further. Starting with the same basic 320- to 1,000-cycle standard variable-frequency a-c system, a full-wave 3-phase bridge rectifier is proposed to provide d-c power. This time a 3-wire 126/252-nominal-voltage system would result. Again the standard a-c generators supply directly, without transformers or direct-voltage regulators, both the variable-frequency a-c and the d-c loads with several of the d-c rectifier outputs connected in parallel.

Outline of Full-Wave 3-Wire System

The block diagram in Fig. 3, while similar to Fig. 1, illustrates the following additional features:

The selenium rectifier is connected

directly without transformer to the a-c generator load bus. The neutral of the wye-connected a-c generator is grounded to the airplane structure which makes the structure the neutral of the 3-wire d-c system. The rectifier is connected as a 3-phase bridge providing full-wave rectification at a nominal voltage of 252, and half-wave 3-phase rectification at a voltage of 126. Referring to Fig. 4, the 126-volt d-c output is supplied from ground to the positive bus. It is to be noted that if 126-volt loads are taken from both sides of the 252-volt rectifier output lines to ground, that the group of loads supplied from the 252-volt negative line to ground would be effectively operating with a grounded positive. To avoid this, no 126-volt loads are shown connected to the 252 volt-negative line and ground. It is obvious, therefore, that the ampere capacity of the rectifier supplying the positive line, which must carry both the 126-volt half-wave and the 252-volt full-wave currents, will be larger than the rectifier sections in the negative line. The characteristics of each of the 126-volt rectifier circuits are identical with the circuit of Fig. 2.

Disadvantages of full-wave system will be similar to those listed under "Disadvantages of System," except the following:

1. Wave-form distortion of the a-c supplies caused by a combination of half-wave and full-wave rectifier loads may prove to be a problem but in any event it should be substantially less severe than with the half-wave loads only. If the loading of the system were entirely on the 252-volt full-wave d-c output, much lower harmonic distortion should be expected. No tests have been made adequately to determine the harmonic percentage of such a full-wave system.

2. Ripple in the full-wave d-c output will be considerably less and ripple in the 126-volt half-wave portion of the d-c load will be the same 18 per cent. (See discussion of 18-per-cent ripple in section entitled "Disadvantages of System.")

Advantages of full-wave system will be similar to those listed under "Disadvantages of System," except as follows:

1. A 126/252-volt d-c power source is obtained from a standard variable-frequency a-c system using the same a-c generator for both a-c and d-c power, and again optimum simplicity is gained through the elimination of transformers, direct-voltage regulators, and reverse-current relays.

2. A 252-volt d-c power source is very desirable for large airplanes, or for systems with large loads of the order of 30 kw or larger in the extremities of the airplane, such as large flight surface control loads. If d-c loads from a single a-c generator are to exceed 50 kw, then to avoid the same distribution problems we now have with high power at 28 volts, a voltage higher than 126, such as 252 volts would be necessary.

3. Ripple voltage would be reduced to 4 per cent on the 252-volt full-wave d-c output.

4. Distribution system saving by the use of 252-volt d-c power on heavy duty loads would be substantial in very large airplanes.

Variations of the Two Systems

Of several possible variations of the two systems described, one notable variation would be a system quite similar to that discussed under "A Combined Variable-Frequency A-C and Full-Wave 126/-252-Volt D-C Generating System," but with a different grounding scheme. Referring to Fig. 4, the 252-volt negative bus would be grounded to the airplane structure, thereby acting as the ground return for both the 126-volt and 252-volt-circuits. This is advantageous as it allows a single wire to be used for circuits of either voltage. It does, however, have the disadvantage of being unable to have the neutral of the 3-phase a-c rectifier grounded, since the neutral is the 126-volt d-c positive feeder. This system is mentioned only as an example of many possible variations and will not be discussed further, as the intent of this paper is not to explore all systems but to suggest a simplified electric system without changes to present standard variable-frequency generating systems.

Conclusions

This preliminary study indicates that a transformerless rectifier combination a-c/d-c aircraft electric system eliminating the need for direct-voltage regulators and reverse-current relays, and providing parallel-operated d-c sources and economical variable-frequency a-c power from the same a-c generators, may be practical. In view of the critical need for simpler, lower cost aircraft electric systems of greater reliability, further study and experimental investigation of the subject systems seems merited.

References

1. Power Waveforms in Aircraft. *Tele-Tech*, New York, N. Y., June 1951, p. 61.

2. Electric Power Characteristics. *Military Aeronautical Specification MIL-E- 7894*, Aug. 14, 1952.

3. Report of Advisory Staff for Aircraft Electric Systems. Wright Air Development Center, Dayton, Ohio, 1951.

4. Higher-Voltage D-C Aircraft Electric Systems, W. L. Berry, J. P. Dallas. *AIEE Transactions* (*Electrical Engineering*), vol. 63, Nov. 1944, pp. 842–48.

5. Save Airplane Weight with D.C., W. L. Berry, J. P. Dallas. *Air Transport*, Aug. 1945.

6. Lightweight Aircraft Transformers, D. S. Stephens. *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 1073–78.

7. Selenium Rectifier Stacks. *Technical Bulletin C-349*, International Selenium Rectifier Corporation, revision Dec. 1, 1951.

8. Basic Considerations in Selection of Electric Systems for Large Aircraft, W. K. Boice, L. G. Leroy, Jr. *AIEE Transactions* (*Electrical Engineering*), vol. 63, June 1944, pp. 279–87.

9. Aircraft Electric Power-Supply System, J. E. Yarmack. *AIEE Transactions*, vol. 62, 1943, pp. 655–58.

◆

Discussion

B. J. Wilson (Naval Research Laboratory, Washington, D. C.): The authors are to be commended on their forthright discussion of a subject whose application to practice may prove expedient in the growth of large and reliable aircraft electric systems. In the event prime mover equipment with constant speed outputs fail to achieve the required degree of reliability at the time when paralleled a-c electric sources become highly desirable, it may be necessary to resort to some substitute measure such as the one described in this paper.

However, several problems arise in connection with the application of the described combined a-c/d-c system. In considering operational reliability, one must include environmental effects. Though admittedly this is not without its difficulties, the obligation remains. Keeping in mind a carefully evolved definition of reliability, the probability of successful performance as intended of a device or system under given environmental conditions and for a given length of time, one is inclined to question the rectifier performance illustrated in the paper as to its contribution to system reliability. Existing high-power dry-disk rectifiers are affected in performance considerably over the broad band of environments to which modern aircraft are subjected. For instance, heating effects whose considerations are classical in the study of dry-disk rectifiers should be analyzed carefully over the expected range of environments. Unequal heating among rectifiers that creates differences in series impedances among the phases can produce polyphase unbalancing of generator operation with its consequent heating effect on other a-c equipment due to negative-sequence current components. In addition, effects from lack of complete magnetomotive cancellation of d-c and/or zero-sequence components may incur further losses and irregularities of terminal conditions.

The feasibility of paralleling to a common d-c bus the rectified output of a-c generators was discussed in an unpublished paper by Cobb, Kershaw, and Erlandson in 1948 and used as early as 1949 in certain critical applications with a resultant weight saving over that of the straight d-c system.[1]

It seems fitting to re-emphasize the point stressed by Mr. Dallas and Mr. Reising on electric system performance limits. After all, the objectives of the electric system are merely one factor in a chain of more general aircraft performance requirements. Equally important are the performance objectives of the electric equipment toward ultimately influencing control surfaces or other aerodynamic and associated guidance or navigational factors. It follows, then, that the most logical way to qualify the generation and distribution system performance is first to assess the electric input limits of equipment that satisfy the output performance limits. One might say it is immaterial what input electric variations are experienced, within reason of course, as long as the output performance objectives and limits of the electric equipment are satisfied.

REFERENCE

1. THE RELATIVE MERITS OF AUXILIARY POWER SYSTEMS—THE ELECTRICAL ASPECT, R. H. Woodall. Second International Aeronautical Conference, Institute of the Aeronautical Sciences, New York, N. Y., May 24–27, 1949, p. 455.

J. P. Dallas and **C. A. Reising, Jr.:** Environmental conditions such as altitude, temperature, humidity, heat, etc., do present many problems which, of course, must be solved before this system can be successful. However, most of the environmental problems that apply to dry-disk rectifiers have been overcome in other items of electric equipment and the situation seems to be similar in nature. The heat to be removed and the maximum temperature allowed in dry-disk rectifiers will require adequate cooling at all altitudes and may need to be derated in a like manner to generators.

Unequal heating among rectifier plates causing differences in series impedances, and therefore undesirable troubles within the generator, are recognized as points to be explored in the pursuit of this proposal. The seriousness of unequal plate impedance will have to be determined by tests. It is anticipated that by careful selection and proper aging of the rectifier plates uniform sections can be successfully built, thus minimizing the phase unbalance.

Mr. Wilson's comment of previous experience in parallelling the d-c output from rectifier is sincerely appreciated. It is gratifying to learn that Mr. Wilson, and it is hoped others, share the views of the authors regarding performance limits and system requirements. In the years since World War II much time and money have been spent trying to perfect a means of parallelling a-c machines to improve system reliability. The system proposed may be one solution to the basic problem of a simple, reliable aircraft electric system.

Considerations Applicable to Automatic Paralleling of Aircraft A-C Generators

M. J. POWELL
ASSOCIATE MEMBER AIEE

E. W. GILOY
ASSOCIATE MEMBER AIEE

Synopsis: Practical and theoretical considerations relative to automatic paralleling of aircraft a-c generators are presented. Different types of a-c generator systems are classified generally as to requirements, and paralleling assumptions are proposed as design criteria. Synchronizing requirements as determined by subtransient and transient disturbance to the system, and paralleling requirements as determined by steady-state load division are presented. Considerations applicable to the development of autoparalleling systems are also discussed.

THE ever increasing complexities of aircraft dictated by operational and functional requirements make it desirable, it not necessary in certain cases, to relieve the airplane crew of manual control of many functions of aircraft operation. Particularly this is considered true for certain control operations of paralleled a-c power systems. Synchronization and paralleling of a-c generators by manual means is not necessarily complex but is a somewhat tedious operation which can require the attention of a crew member when other duties are more demanding of his time. Also, experience to date has demonstrated that the paralleling operation may be inadvertently performed improperly by the operator, producing undesirable transients and improper operation of the system.

Automatic paralleling of a-c generators is not in itself new, but in considering its possible application to aircraft it is believed this new environment does warrant a reconsideration of objectives and requirements. It is desired, therefore, to present a few thoughts considered pertinent to the subject.

Classification of Aircraft A-C Systems

For purposes of discussion and consideration of the subject it is believed advisable to classify a-c power systems as follows:

1. Stiff system: Generators in this case may be coupled directly to the prime mover and, if main engine driven, will in general be wide-frequency range machines.

2. Semisoft system: Generators may be driven by a mechanical-hydraulic differential type of transmission through an overrunning clutch restricting the transmission of torque between generator and drive to a single direction.

3. Soft system: Generators may be coupled to their basic source of power through an intermediate device, such as an

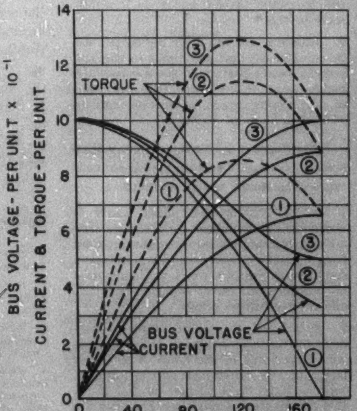


Fig. 1. Subtransient current, voltage, and torque at instant of paralleling versus phase angle between incoming generator and bus. Number in circle indicates number of generators connected to bus prior to paralleling. Calculations for curves were based on generator $X_d''=15$ per cent. (Information courtesy of Westinghouse Electric Corporation)

Paper 54-312, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted February 15, 1954; made available for printing May 17, 1954.

M. J. POWELL and E. W. GILOY are with The Glenn L. Martin Company, Baltimore, Md.

The authors take this opportunity to express their appreciation to members of The Glenn L. Martin Company Engineering Division who cooperated in the preparation of this paper.

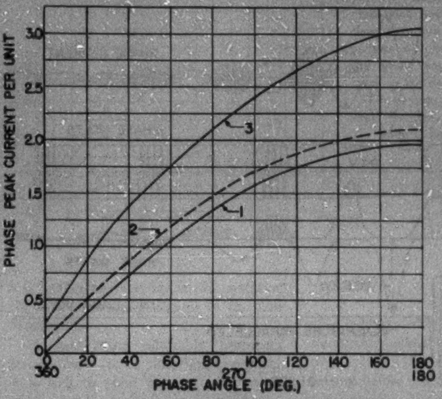


Fig. 2. Transient current as a function of synchronizing angle. One generator on bus with the following load conditions prior to paralleling: 1. zero bus load; 2. 40-kva 0.75-power factor bus load; 3. 40-kva 0.75-power factor bus load with 30-kw load on the incoming generator. Generators operating at 115 volts line to ground and 400 cycles per second

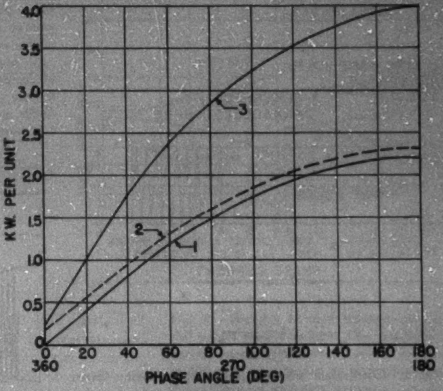


Fig. 3. Transient power as a function of synchronizing angle. One generator on bus with load conditions as in Fig. 2

air turbine motor driven by bleed air from the airplane's main engine. This system will also probably include overrunning clutches.

System 1 may not be capable of parallel operation owing to peculiarities of prime mover speed operation as in the case of an airplane main engine. Should the prime mover be speed-regulated, the stiffness of coupling between generator and prime mover might in general require a paralleling accuracy exceeding that considered satisfactory for other systems. System 3 is usually more limited in regard to per cent of rated torque that can be momentarily delivered to the generator and, therefore, may be more lenient as to paralleling accuracies needed. System 2 is perhaps average as compared to systems 1 and 3 in regard to paralleling accuracy requirements.

In general, the synchronizing requirements presented in this paper are more directly applicable to the second system. The test data for the curves presented were obtained on a system where two constant-speed drives, regulated by flyball-type governors in tandem with servomotor control, were utilized to drive two 40-kva 0.75-power factor 115/200-volt 400-cycle a-c generators controlled by carbon pile-type voltage regulators.

Classification of Paralleling Systems

Again for purposes of consideration and discussion, it is well to classify paralleling systems according to their degree of automaticity:

1. Automatic system: Generators automatically synchronized and connected to the bus.
2. Semiautomatic systems:
   a. Generators manually synchronized and automatically connected (paralleled) to the bus.
   b. Generators automatically synchronized and manually connected to the bus.
3. Manual system: Generators manually synchronized and connected to the bus.

Paralleling Assumptions

As design criteria it is proposed that:

1. It is desirable to limit the maximum transient disturbance, occuring approximately 10 to 20 milliseconds after paralleling, to generator full-load current, with transient voltage at all times as defined by the military aircraft power specification *MIL-E-7894*.[1]

2. The initial performance of a properly operated system should maintain the difference in load, both real and reactive, between the generator with the greatest load and the generator with the smallest load within 10 per cent of a single generator rating, watts and vars. No service adjustment of voltage- or frequency-control equipment should be required until the

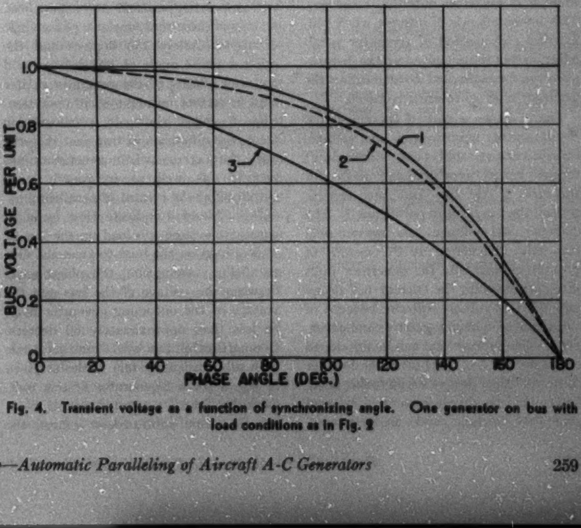


Fig. 4. Transient voltage as a function of synchronizing angle. One generator on bus with load conditions as in Fig. 2

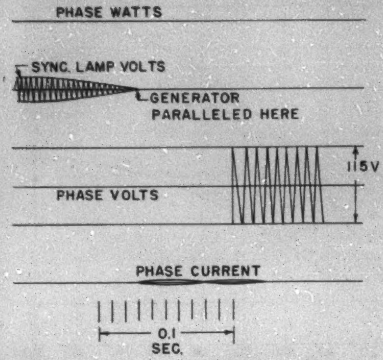


Fig. 5. Oscillogram showing paralleling transient. Generators paralleled at a synchronizing angle of approximately zero degrees when operating at zero bus load, 115 volts rms line to ground and 400 cycles per second

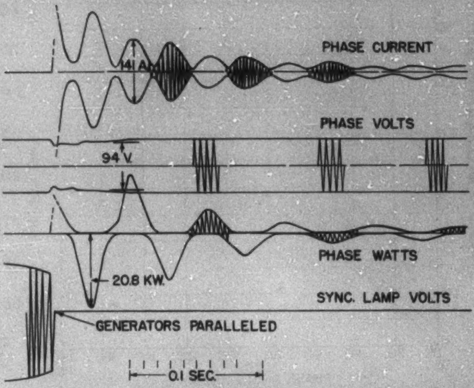


Fig. 6. Oscillogram showing paralleling transient. Generators paralleled at a synchronizing angle of approximately 115 degrees. Bus at zero load, 115 volts line to ground and 400 cycles per second. Incoming generator 5 cycles per second high. Voltages and currents are rms values calibrated from peak to peak of their respective trace. Kilowatt value represents average of peak power as would be indicated by a visual wattmeter measuring continuous power corresponding to this peak value

difference in load between machines has reached a value of 20 per cent.

Synchronizing Requirements Determined by Subtransient and Transient Disturbances to the System

When an a-c generator is paralleled with the bus, the bus voltage transient, inrush current, and torque developed at the instant of breaker closing are determined by the phase angle between the voltage of the oncoming machine and that of the bus, the voltage difference between them, and the internal impedance of the generators. These values are also affected by the frequency error, and the number of machines initially on the bus. This information is of interest when considering any method of generator paralleling, so that the magnitude of the disturbance can be estimated to determine the accuracy of synchronizing required.

Fig. 1 shows a plot of the calculated instantaneous subtransient bus voltage, inrush current, and peak subtransient torque versus synchronizing angle for the cases of one, two, and three generators on the bus prior to paralleling.[2] The currents and voltages are symmetrical rms values occurring at the instant of breaker closing. As the generator pulls into synchronism, the current will decay and the bus voltage will rise because of the action of voltage regulators and governors. The subtransient torque developed occurs within 1/2 cycle after the breaker closes and then decays rapidly, this being determined by the decrement of internal generated voltage and the changing torque angle as the generators pull into step.

Figs. 2, 3, and 4 were obtained from tests on a system using two 40-kva a-c generators driven by mechanical-hydraulic differential-type drives. They show peak transient current, peak transient power, and the transient voltage plotted against the phase angle between the bus voltage and the voltage of the oncoming generator. These values were taken from oscillographic records approximately 10 to 20 milliseconds after the instant of paralleling. A generator exciter polarity reversal was experienced on two of these runs, one at a phase angle of approximately 130 degrees and the other a phase angle of 180 degrees, and required "flashing" the generator exciter field to restore normal system operation. Figs. 5 and 6 show, by oscillographic trace, the difference in transient disturbance that occurs when generators are properly as well as improperly synchronized at the instant of paralleling.

Figs. 2 and 4 indicate that for a 2-generator system (no load on the bus or 40-kva load on the bus) to meet the first paralleling assumption, the phase angle between the voltage of the bus and the voltage of the oncoming generator must be less than approximately 60 degrees. Comparison of this synchronizing angle with subtransient system values as given in Fig. 1 for a 2-generator system indicates that for a 4-generator system, holding the same subtransient values, the synchronizing angle will probably need to be held to a maximum of 30 to 35 degrees. To allow for such multiplicity of generators in an aircraft power system, it is recommended that generators be synchronized to a phase angle not exceeding approximately 30 degrees at the instant of paralleling. This value of synchronizing angle is based on the assumption that the bus and generator voltages are equal in magnitude. However, a difference in scalar magnitude of the voltages of even 5 or 6 volts, when based on a phase voltage of 115, will be small compared to voltage differences due to a phase-angle error, hence its effect on the transient disturbance will be small. The frequency error within reasonable limits at the time of paralleling is also not too important if overrunning clutches are used, but it does mean that a fast machine when paralleled will momentarily attempt to take all of the load. The frequency error also determines the time available for sensing the phase angle and closing the necessary control gear to connect the oncoming generator and bus in parallel. Regarding this, Fig. 7 shows that for a phase-angle difference of 30 degrees and 50 milliseconds operating time (assumed as the time required for necessary control equipment operation) the bus and generator frequency difference cannot exceed 3 cycles. In view of this, it is apparent that the frequency error at the instant of paralleling must not exceed approximately 3 cycles and

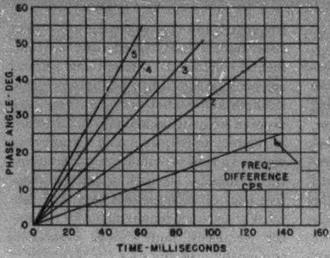


Fig. 7 (left). Synchronizing angle versus time that two rotating vectors will be within the given angle, i.e., from −10 degrees to +10 degrees, etc., if their frequency difference is known

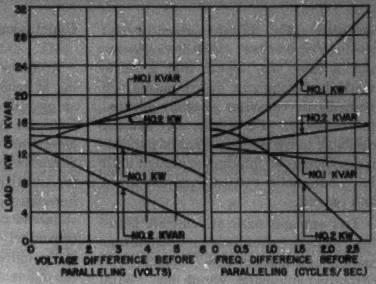


Fig. 8. Strength of load division circuits. All values read 5 minutes after paralleling. For voltage curve, frequency was held constant at 400 cycles per second on both generators, and voltage of incoming generator was decreased while voltage of no. 1 was held at 115 volts. For frequency curve, voltage was held constant at 115 volts and frequency of incoming generator was decreased from 400 cycles per second. In all cases values were set and load applied before paralleling

it is further desirable under this condition that the oncoming generator be the slow machine.

In considering allowable frequency errors, phase angles, and time to parallel, it is well to point out that the values given previously are based on the assumption that the generator drive speed governing is steady and unmodulated. Experience to date has indicated that this may not always be true. If generator drive speed modulation is present (defined as a "spasmodic frequency jump" where the frequency will change from 3 to 5 cycles almost instantaneously and be damped out in approximately 1/2 second) it is reasonable to assume that under such operating conditions the average difference in bus and oncoming generator frequency may well need to be closer than 3 cycles to allow for some disturbing effect of generator frequency variations. It is advisable, therefore, that the instantaneous as well as the average regulation accuracy of the generator drive be specified for the proper application of automatic synchronizing methods.

Paralleling Requirements Determined by Steady-State Load Division

In view of the need and the desirability for using automatic paralleling control, it follows, if full benefits are to be obtained, that any such system should not require manual adjustments after paralleling. This demands that the initial frequency and voltage settings, at the instant of paralleling, be held within limits that allow the generators, controlled by the load division circuits,[3] to share the bus load within the accuracy defined by the second paralleling assumption.

Fig. 8 shows the load division of a 2-generator system presently in use. This curve presents the system reactive load division when the generator frequencies are held constant to 400 cycles per second and the voltage setting before paralleling of one generator is varied. This curve also indicates the real load division of the system when the voltages are held constant to 115 volts and the frequency setting before paralleling of one generator is varied. To obtain the desired load division, within 3 kw and approximately 3 kilovars in this case, Fig. 8 shows that the voltages cannot differ by more than 0.75 volt and the frequencies cannot differ by more than 0.75 cycle per second. These values of allowable voltage and frequency errors are of course far short of the differences between generators that can exist within the regulation range of the present system voltage and frequency regulation equipment and, therefore, proper real and reactive load division cannot be obtained at all times without occasional manual monitoring of the system. The strength of the system's real and reactive load division circuits is obviously too weak for automatic paralleling control.

Fig. 8 also indicates that the real and the reactive load-sharing circuits are not entirely independent of each other and need improvement in this regard as well. The real load-sharing sensing in particular appears to be too sensitive to vars' unbalance.

To obtain proper real and reactive load sharing will require an over-all consideration of system load-sharing control strength and the regulation range of system voltage and frequency control equipment. Experience has indicated that stability considerations, in regard to load sharing circuit strength, may well require the development of system voltage and frequency regulation equipment of greater accuracy over all load, speed, and environmental conditions than is presently used.

Development of Automatic Paralleling Systems

In considering general methods of obtaining automatic paralleling for aircraft a-c systems, two different possible approaches stand out:

1. One system might include necessary automatic adjustment of generator frequency and voltage by some suitable servo control system in conjunction with other synchronizing control to bring generator frequency, phase, and voltage into agreement with the corresponding values of the bus or other generator. When properly synchronized, the oncoming generator may then by other suitable means be automatically closed on to the bus or manually selected by the operator as airplane operating conditions dictate. In parallel operation, suitable real and reactive load division circuits will automatically provide for proper load sharing.

2. Another system might utilize generator drives, including overrunning clutches, and voltage regulators so designed and developed that for all speed, load, and environmental conditions, generator frequency and voltage deviations from that of the bus or other generators will not preclude automatic load sharing when generators are paralleled. With such generator drives and regulators, the need for a special synchronizing servo control will probably be eliminated. Some device will be necessary, however, to sense automatically generator and bus voltage phase relationships and to close the oncoming generator to the bus within the previously discussed phase limit of 30 degrees.

In the development of automatic paralleling systems, it may be necessary to approach the final product on a step-by-step basis. Time will undoubtedly be required to develop drives and voltage regulators of the desired narrow regulation range as needed for method 2 and, considering the different presently used methods of constant speed drive governing, it does not appear that a single method of automatic synchronizing will be easily developed or generally adapted for method 1. These facts may make it desirable to use a semiautomatic paralleling system as an interium expedient until fully automatic systems are developed and proved.

Two semiautomatic systems are possible, namely, automatic synchronizing-manual paralleling[4] and manual synchronizing-automatic paralleling. With the former system, one is again confronted with differences in drive governing methods in use today. One method utilizes an electromechanical type of control in which governing is obtained by the balancing of spring and solenoid forces. Trimming changes in governor regulation are obtained by varying the over-all electric signal strength supplied to the governor solenoid. A second method of drive governing utilizes a mechanical flyball type of control, and governor trimming is obtained by varying the over-all spring biasing of the governor through action of a small electric motor and cam assembly.

The response of governor trimming and the control equipment which is used are not the same for these two methods of governing. A synchronizing control that may be applicable to one may not be adequate for the other.

For the semiautomatic system, namely, manual synchronizing-automatic paralleling, it is possible to develop a device to sense when generator and bus voltage phase relationships are proper, and to close automatically an incoming generator to the bus without involving or affecting the different methods of drive governing. In using this method of paralleling it will only be necessary for the system operator to vary the frequency of the incoming generator until it closely approaches or passes through the frequency setting of the bus. At this time, a phase-sensing device can automatically connect the incoming generator to the bus at the desired synchronizing angle. In view of present-day real load equalizing circuit strength, some readjustment of frequency setting will perhaps be required after paralleling to equalize generator loads, but the actual operation to parallel requires little or no accuracy and could well be performed blindfolded. It is suggested, therefore, that the manual synchronizing-automatic parallel method of semiautomatic systems be considered as the first step toward a completely automatic system.

Conclusions

Accuracy limits for automatic paralleling systems should be determined from an over-all consideration of the following factors: 1. The transient disturbance, determined largely by the phase angle between the voltage of the oncoming generator and that of the bus, and the internal impedance of the generators; 2. the stabilized operation of the system while in parallel, determined by the difference in the induced voltage of the generators, the difference in frequency setting before paralleling, the strength of the reactive load division circuit, and the strength of the real load division circuit.

This preliminary investigation based primarily on aircraft generators of a 40-kva size, and covering the synchronizing requirements necessary to meet the transient requirement as defined by the paralleling assumptions for a system covering up to four generators, indicates that the phase angle between generator voltages must be less than 30 degrees at the instant of paralleling. The frequency difference between machines does not appear critical when using overrunning clutches but from practical considerations it seems that 3 cycles or less are required if sufficient time is to be allowed for control action. The difference in machine-induced voltage at the time of paralleling also does not appear to be critical within practical limits, and machine voltages may be any value within the nominal regulation accuracy of 2 or 3 per cent obtained with present-day voltage regulators.

The frequency and voltage differences that may be permissible between generators in regard to the initial paralleling transient may well be too severe to meet the objective of a fully automatic (i.e., no-manual monitoring required) system as defined by the paralleling assumption for stabilized parallel operation. The initial voltage and frequency differences that can be tolerated and that still provide for desired generator parallel operation will of necessity be determined by the strength of the respective generator system's real and reactive load division circuits.

If the full advantages of automatic paralleling are to be obtained, it is necessary that future controls for a-c generator systems do not need or include a remote means of generator frequency or voltage adjustment. Therefore the real and reactive load division circuits must be capable of providing the desired load sharing over the complete regulation range of the particular generator and drive frequency, and voltage control equipment covering all load, speed, and environmental conditions. Present-day systems are weak in regard to load division signal strength and cannot provide proper load division without occasional readjustment of the remote-frequency and voltage controls. In view of stability considerations, the permissible strength of real and reactive load division circuits may well dictate the accuracy requirements for generator frequency and voltage regulation equipment. In any case the objective is defined and the solution will require consideration of all factors involved.

The development of paralleling systems may well follow one or more of these methods:

METHOD A, SEMIAUTOMATIC

1. Remote control of generator frequency and voltage to allow for manual synchronizing of generator and bus.
2. Development of a device to sense when generator and bus phasing is proper, within 30 degrees and to close automatically the oncoming generator to the bus.
3. Development of equalizing circuits stronger than those presently used.

METHOD B, AUTOMATIC

1. Development of nessessary servo control for automatic adjustment of generator frequency and voltage coupled with other suitable control to synchronize generator frequency and phase closely to that of the bus.
2. Development of an automatic phase-sensing device as in method *A*, item 2.
3. Development of real and reactive load division circuits of sufficient strength.

METHOD C, AUTOMATIC

1. Development of generator drives and voltage regulators of sufficiently narrow regulation range that proper load division can be obtained in parallel operation, utilizing necessary load division circuits, over all speed, load, and environmental conditions.
2. Development of an automatic phase-sensing device as in method *A*, item 2.
3. Development of real and reactive load division circuits of sufficient strength.

Method *A* is recommended as the first interim step in the development of automatic systems, and method *C* is considered to be the most desired system.

References

1. AIRCRAFT CHARACTERISTICS OF ELECTRIC POWER. *MIL-E-7894*, Bureau of Aeronautics, Washington, D. C., April 28, 1952.

2. Transient Electric Torque of Turbine Generators During Short Circuits and Synchronizing, H. S. Kirschbaum. *AIEE Transactions* (*Electrical Engineering*), vol. 64, Feb., 1945, pp. 65–70.

3. POWER EQUALIZER SYSTEMS FOR AIRCRAFT

Alternators, J. A. Granath, A. K. Hawkes. *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 209–217.

4. Automatic Synchronizing for Aircraft Alternators, F. B. McCarty, G. W. Hills. *AIEE Transactions*, vol. 71, pt. II, 1952 (Jan. 1953 section), pp. 454–58.

◆

Discussion

Russell W. Stineman (Boeing Airplane Company, Seattle, Wash.): To the authors' "Classification of Paralleling Systems," it is suggested that a fourth category be added, "Random System," generators are manually connected to the bus with no prior attempt to synchronize. There is some evidence that random paralleling of aircraft generators may be feasible under certain conditions and in any event should be considered along with the other systems.

Subtransient currents and torques that occur during paralleling are of such short duration in typical aircraft generators that they may be ignored in arriving at design criteria. It would therefore appear that the authors' first paralleling assumption could be safely modified to allow transient currents within the generator overload rating, i.e., two per unit current for 5 seconds.

P. F. Boggess (Westinghouse Electric Corporation, Lima, Ohio): The authors have clearly presented the practical and theoretical problems of paralleling and synchronizing aircraft a-c generators. They have classified the paralleling of a-c systems into three categories and have defined each system as to automatic systems, semiautomatic systems, and manual systems.

This discussion is to clarify the definition of synchronizing versus paralleling, and to point out that automatic paralleling does not necessarily require synchronizing before paralleling. Automatic synchronizing may be defined as the automatic adjustment of frequency, voltage, and phase relationship of the incoming generator, so that its output is maintained identical, within limits, with the bus. Thus, paralleling can occur at any time.

Automatic paralleling may be defined as the act of closing the a-c generator breaker paralleling the incoming generator with the bus when frequency, voltage, and phase relationships of the incoming generator are within limits that do not cause severe system transients.

Present specifications for constant-frequency power specify 380/420 cycles per second as system frequency. Synchronization would be required if the complete range of 380 to 420 cycles per second were used in parallel systems. However, with currently available constant-speed drive control and voltage regulator control, frequency and voltage are each held within limits that permit satisfactory paralleling. A relay that senses when the phase relationships of the incoming generator and the bus are satisfactory for the specified transient limit is then all that is needed for paralleling control.

Some consideration is being given to providing automatic synchronization, but such systems in the past have been heavy and complicated (see reference 4 of the paper). Where constant-speed drives are used that control frequency within ±2 cycles per second, no special provision for paralleling is required except the autoparalleling relay.[1] Nearly all parallel systems currently planned are paralleled with an autoparalleling relay.

Reference

1. New Developments in Differential-Type Hydraulic Transmissions and Controls, R. H. Eisengrein. *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 423–29.

M. J. Powell and E. W. Giloy: The authors express their appreciation for the interest in and comments on this paper. In regard to the fourth category "Random System" suggested by Mr. Stineman, it is important to consider the disadvantages mentioned in the paper, namely, exciter reversals and high subtransient and transient torques experienced. If random paralleling is to be employed, full evaluation must be made of these items as well as system electric parameter transients. A random paralleling system would also require utilization of overrunning clutches to limit transmission of torque from the generator to its prime mover.

In regard to Mr. Boggess's comments, the authors agree with and appreciate his clarification of the definitions given in the paper.

A Method of Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems

S. C. CALDWELL
ASSOCIATE MEMBER AIEE

L. E. JENSEN
ASSOCIATE MEMBER AIEE

Synopsis: When a fault occurs on an aircraft a-c system with voltage-regulated alternators, the short-circuit current goes through a transient period depending on the alternator-regulator characteristic. This transient period has been found to vary from several milliseconds to perhaps 300 milliseconds and, consequently, may have a considerable effect on fuse or current-limiter clearing times. This paper presents a method of calculating current-limiter or fuse-clearing times which includes the effect of this transient period. Results of calculations from this method are compared with laboratory tests demonstrating that the method has a reasonable degree of accuracy.

AIRCRAFT a-c distribution systems rely almost entirely on thermal devices such as current limiters, fuses, and circuit breakers for fault protection. Application and co-ordination problems make necessary a knowledge of the clearing times of the devices concerned under fault and other system transients. Manufacturers of these devices provide current-time curves based upon a constant-current supply. If actual clearing times in system operation are to be determined, the transients in the currents through the devices must be accounted for. Where clearing times are short enough so that the transient duration is an appreciable portion of the clearing time, the effect of the transient can be large.

This paper presents a method for the determination of clearing times together with test results for transients typical of aircraft a-c systems. This method is applied to limiters widely used in existing aircraft distribution systems. It permits calculation of clearing times for the individual lines as well as total clearing times for faults with multiple feeds. Comparisons with laboratory test results are presented.

Analysis

In thermal devices, the heating is a function of the square of the current, the instantaneous resistance of the element, and the duration of the current flow through the element. If heat losses due to radiation, conduction, and convection are neglected, the following equation holds for any element of the fuse wire

$$dq=k_1i^2rdt$$

where

Paper 54-307, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted July 1, 1953; made available for printing April 23, 1954.

S. C. Caldwell and L. E. Jensen are with the General Electric Company, Schenectady, N. Y.

The authors wish to acknowledge the suggestions and assistance of M. N. Halberg of the General Electric Company, Schenectady, N. Y.

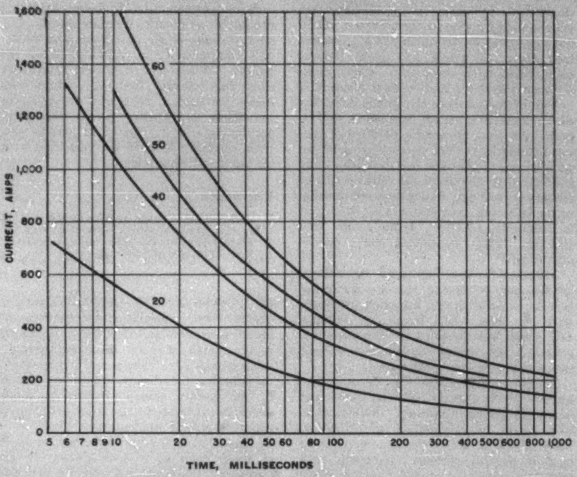


Fig. 1. Current limiter current-time curves

q = energy stored in the wire element
i = instantaneous current in the element
r = resistance of the element as a function of the temperature θ
k_1 = constant depending on units used
t = time

Assuming that the heat storage coefficient of the wire element is independent of temperature and the resistance is linear with respect to temperature

$$dq=k_2d\theta=i^2r_0[(1+\alpha(\theta-\theta_0)]dt$$

where

k_2 = constant dependent upon the heat storage coefficient of the element and the units used
θ_0 = reference temperature at which element resistance is measured
r_0 = resistance of the element at reference temperature
α = temperature coefficient of resistance of the element

Integration of this equation results in the following

$$\int_{t_1}^{t_2} i^2dt=\frac{k_2}{r_0\alpha}\log_e\left[\frac{1+\alpha(\theta_2-\theta_0)}{1+\alpha(\theta_1-\theta_0)}\right] \qquad (1)$$

where

t_1 = initial time
t_2 = final time
θ_1 = initial temperature of the element
θ_2 = final temperature of the element

It is evident from this last equation that the temperature rise of the element is dependent only upon the time integral of the square of the current. Thus, over the time range in which the assumptions hold, the shape of the current transient has no effect on the total temperature as long as the time integral of the square of the current remains constant. This suggests that the effects of the transients within the duration of time over which this equation holds may be accounted for simply by taking the time integral of the current squared, and comparing this quantity with the time integral of current squared, derived from the limiter constant-current curves furnished by the manufacturer.

As a means of determining the range over which equation 1 is valid for limiters, the curves of the time integral of current squared were plotted against time in Fig. 2 for the constant-current data of Fig. 1. A curve for each of the limiters tested is shown. If equation 1 is to hold, these curves must be constant with respect to time. It is seen that this condition is approximated to a reasonable degree of accuracy up to about 0.3 second for the limiters. The slight rise noticeable at short times is probably accounted for by the fact that, as clearing times become faster, the arcing time after melting of the limiter element may become an appreciable part of the total clearing time. Thus, the time integral of current squared may appear larger than is indicated by equation 1 which does not include the effect of arcing. The rapid rise of the curves beyond 0.3 second is attributed to the supposition that, since heat losses are an integral function of time, they are negligible for short times but become large as time increases.

It would appear, then, that clearing times for aircraft limiters may be obtained with reasonable precision, by comparison of time integrals of current squared for transient conditions with the same integrals for the constant current data available for the limiters, since fault transients in aircraft systems generally last for 0.3 second or less. This compari-

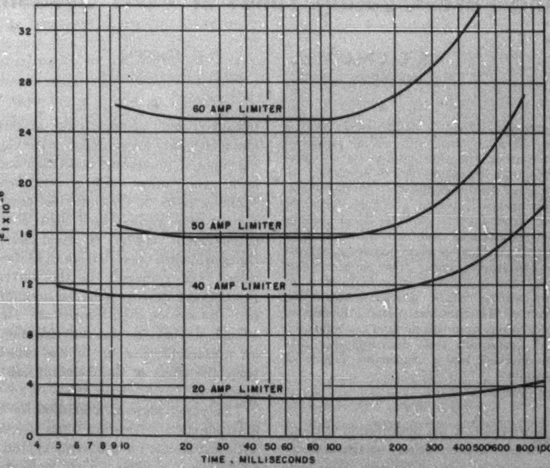


Fig. 2. Time integral curves for aircraft limiters

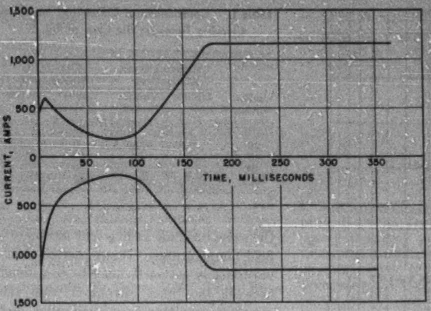


Fig. 3. Envelope of short-circuit current for single-phase line-to-ground fault

son is enhanced by the fact that the accuracy of the comparisons depends primarily upon the difference in the heat losses resulting during the transient as compared to those resulting from a constant current over the same period.

Faults With Single Feed

In the case of faults with single feed, only one transient in the limiter current is involved, therefore it becomes convenient to obtain an equivalent constant current versus time curve for the transient, so that direct comparison may be made with current-time curves for the limiters. It has been assumed on the basis of the preceding analysis that, when

$$\int \frac{dq}{r} = \int_0^t i_f^2 dt = \int_0^t i_c^2 dt = i_c^2 t \qquad (2)$$

where

i_f = instantaneous rms value of transient-fault current
i_c = constant-fault current
q = energy stored in element
r = resistance of limiter element

the limiter will clear.

Thus a constant current

$$i_c = \sqrt{\frac{1}{t}\int_0^t i_f^2 dt} \qquad (3)$$

will clear the limiter in the given time. Therefore, if this value of current is plotted versus time, the intersection of the resulting curve and the current-time curve for the limiter will give the clearing time.

Fig. 3 shows an envelope of a fault-current transient for a 400-cycle aircraft alternator-regulator system. This particular transient shape was chosen since it represents about as great a change, over as long a period, as any produced by aircraft systems familiar to the authors. The ordinate of Fig. 3 is scaled in rms amperes so that at any instant of time t, the rms-current (i_f of equations 2 and 3), may be found directly. After obtaining the rms current as a function of time the integral of equation 3 may be evaluated by squaring the rms current at each time; plotting a curve of i_f^2 versus time and then finding the areas under this curve at succeeding times during the transient period. The effective or virtual current i_c of this transient is plotted as a function of time in Fig. 4, and it is seen to cross the 60-ampere limiter characteristic and predict a clearing time of 150 milliseconds. The tested clearing time for this particular case was found by oscillograph to be 168 milliseconds. Fig. 6 shows comparisons of calculated and test values for this type of fault for several different sizes of limiters. It should be noted that for short clearing times the d-c offset in the alternator transient must be included in the calculation of the equivalent constant current i_c if reasonable agreement is to be expected. This procedure is easily adapted to the case of a fault with multiple feeds where each feed has the same sized limiter and the same current.

The derivation of this method has been straightforward and the method has been used with success by others in the fields of industrial and utility systems.[1-4] The verification herein of its validity for aircraft systems leads the authors to believe that the assumptions used in the derivation are justified and that the derivation may be extended with validity to cases where a fault has multiple feeds.

Faults With Multiple Feeds

In most aircraft distribution systems more than one line per phase connects the alternator, or main bus, to a particular load bus. The arrangement of these lines is such that faults are not likely to occur on all the lines of one phase at the same time. Referring to Fig. 5 with four lines per phase between the main bus and load bus, a fault at the location shown will be fed through both limiters 1 and 2. If the fault is close to limiter 1, and the four lines have equal impedance, three-fourths of the fault current will flow through limiter 1 and one-fourths through limiter 2. Assuming that the changing limiter resistance has no effect on the current magnitude or distribution, limiter 1 will clear first, then all the current will flow through limiter 2 until it clears.[5]

In Fig. 5 it can be seen that by taking

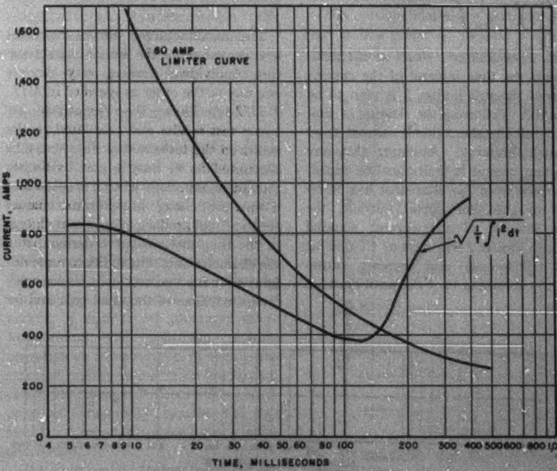


Fig. 4. Fault with single feed

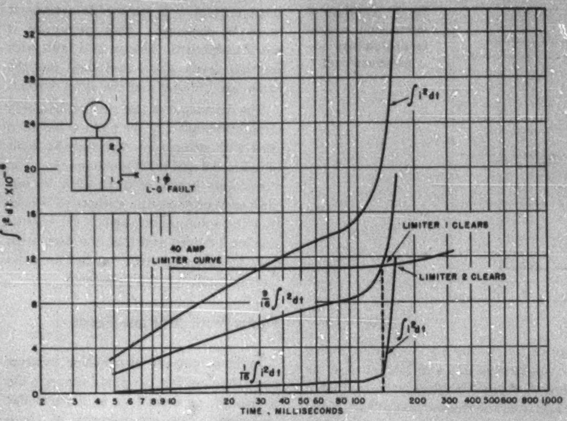


Fig. 5. Fault with multiple feed

the time integral of the current squared through limiter 1 a curve of

$$\int\left(\frac{3}{4}i\right)^2 dt = \frac{9}{16}\int i^2 dt$$

will be obtained.

The intersection of this curve with the 40-ampere-limiter curve will give the clearing time (134 milliseconds) for limiter 1. During this time limiter 2 has been carrying one-fourth of the total current so that the time integral of the current squared through it is

$$\int\left(\frac{1}{4}i\right)^2 dt = \frac{1}{16}\int i^2 dt$$

Thus, when limiter 1 clears at 134 milliseconds the time integral of the current squared through limiter 2 is seen to be $1.6 \times 10.^6$ Following the clearing of limiter 1, all the short-circuit current flows through limiter 2. Assuming that any resultant change in fault-plus-line impedance seen from the alternator has no effect on the short-circuit current, the time integral of the current squared through limiter 2 will start at 1.6×10^6 at 134 milliseconds and increase subsequently according to $\int i^2 dt$ until limiter 2 clears.

By following these steps, clearing times can be found for faults for any number of feeds by taking the new distribution of current following each successive clearing and adding the value of the time integral of the current squared at the time of change of current distribution to the new time integral after redistribution until the clearing time is reached. Table I shows comparisons of test times obtained from oscillographs to calculated times for the case illustrated in Fig. 5.

Discussion

The manufacturers' curves represent average values for the limiter characteristics. Individual limiters may deviate one way or the other as specified in *MIL-F-5372A*.[6] Thus, the deviations between test results and calculated results shown in the table and in Fig. 6 may be accounted for at least in part by limiter tolerances, especially when it is noted that a 3-per-cent change in the limiter characteristic could produce a 6-per-cent change in the time integral curve derived from the characteristic. In addition, some differences in fault current exist between different machines of the same type and for the same machine at different alternator-field temperatures. Hence, actual clearing times in a system must lie in a band between the times calculated from the extremes resulting from generator and limiter tolerances and variations over the expected temperature ranges. This spread in clearing times may prove to be quite large in some cases. In addition, where long, high-impedance lines are involved in single-phase faults, regulator action may be such as to reduce fault current substantially below those characteristic of low-impedance faults, so this must be factored in where necessary.

Table I. Comparison of Calculated and Test Clearing Times for the Case of Fig. 5

Limiter Rating	Calculated Time, Milliseconds		Test Time, Milliseconds	
	Limiter 1	Limiter 2	Limiter 1	Limiter 2
20	9.1	17.7	8.5	17
40	134	155	117	128
40	134	155	115	126
60	172	200	162	185

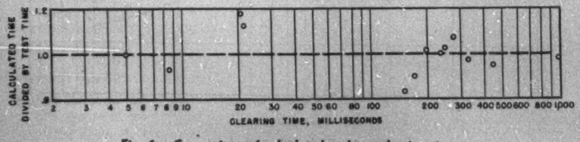


Fig. 6. Comparison of calculated and test-clearing times

At the present stage of development, there are wide variations in shapes of fault-current transients in aircraft a-c alternator-regulator systems. These variations result from significant differences in time constants, regulators, and exciters, and in going from one rating to another and from one manufacturer to another. It would appear desirable from the application standpoint if typical short-circuit current transients could be derived and used universally. With the variety of transients possible, this procedure could lead to considerable error in calculation of clearing times. Thus, to apply the method of this paper, it will be necessary to have available envelopes of fault-current transients for the type and rating of the alternator-regulator system under consideration.

Conclusions

1. A mathematical development has been presented which justifies the use of virtual current curves for calculating limiter clearing times for singly fed faults.

2. The analytical development has been extended to permit calculation of limiter clearing times for multiple-fed faults.

3. Comparisons of test and calculated results have been presented. By showing close agreement between calculations and tests for aircraft limiters, they indicate that the assumptions used to derive the methods of calculation are satisfactory.

4. Though it has not been shown here, it is believed that these methods may be adapted successfully to other thermal devices designed for fault protection in aircraft systems.

References

1. The Application and Standardization of High Rupturing Capacity Current-Limiting Fuses, J. W. Gibson. *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 126–32.

2. The Testing of Electric Fuses With the Cathode Ray Oscillograph, J. A. M. van Liempt, J. A. de Vriend. *Philips Technical Review*, Eindhoven, Holland, vol. 4, no. 4, 1939.

3. Performance Criteria for Current Limiting Power Fuses—II, E. W. Boehne. *AIEE Transactions*, vol. 65, 1946, pp. 1034–45.

4. Transient Performance of Electric Power Systems (book), Reinhold Rudenberg. McGraw-Hill Book Company, Inc., New York, N. Y., 1950, p. 442.

5. Change in Resistance of Aircraft Current Limiters and Its Effect on Current Division in Networks, H. Oman. *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 231–38.

6. Fuse and Closed Link, Aircraft. *Joint Army, Navy and Airforce Specification MIL-F-5372A*, Washington, D.C., Apr. 18, 1952.

◆

Discussion

I. Matthysse (Burndy Engineering Company, Norwalk, Conn.): The authors have clearly described a practical method for using the manufacturer's limiter or fuse-clearing curves to calculate the clearing time, when the fault current is not steady. Although the method of using values of I^2t has been used by others in limiter or fuse calculations, the authors have offered convincing proof of the method, and have outlined a very simple procedure to apply it to practical problems.

S. C. Caldwell and **L. E. Jensen**: The discussion by Mr. Matthysse is greatly appreciated. As noted by the discusser, and in the text of the paper, the principle of using I^2t for singly fed faults is indeed venerable. However, it was felt that extrapolation of the basic ideas to multiple fed faults together with experimental verification of the results for aircraft devices would prove of value to the industry.

High-Speed Rapid Transit Equipment

S. J. VOUCH
ASSOCIATE MEMBER AIEE

Synopsis: Traffic congestion resulting from the increased use of vehicles on our city streets has reached the point where drastic measures are needed to permit freer movement of people within the metropolitan area. One of the brightest hopes for relief is high-speed rapid transit operating on private rights of way. This paper describes the performance which can be obtained with a conventional rail rapid transit system utilizing recent advances in propulsion equipment.

Speed and Convenience Important Factors

IT seems hardly necessary to dwell on the desperate need for improvement in the facilities for moving people within our cities. The gravity of the situation is painfully obvious to every individual whether he travels by private vehicle, by mass transit, or on foot. An analysis of the reasons for this rapidly mushrooming condition serves to point the way to a solution. The greatly increased use of automobiles, brought about largely because of the potential speed and convenience afforded by this means of transportation, is mainly responsible for our worsening traffic situation. The automobile driver is willing to pay the relatively higher cost of personalized transportation as long as it gives him reduced travel time and more convenience than other forms of transportation.

It is reasonable to assume that if the same factors which make the automobile so desirable could be equaled by mass transit, more riders would prefer this transportation medium. The result would be freer movement of people within the city and less need for multiple expressway construction. Experience has shown the fallacy of expecting superhighways alone to provide the relief so sorely needed. By their very nature they invite more motorists to drive into already congested areas. As these swollen traffic streams leave the expressways and enter the business area, they only compound the surface confusion already existing.

Rapid Transit on Expressways

If a mass transit system is to provide the incentives needed to make people prefer it to the private automobile, it must be segregated from other traffic and designed to operate at speeds which will produce less travel time. If this is to be achieved, running speeds must be substantially higher than permitted for automobiles because of the necessity of making a number of passenger stops. Maintaining such schedules requires operation on private rights of way, either elevated, underground, or on grade-separated expressways. From the standpoint of community benefits the last mentioned represents the most attractive alternative.

Two outstanding factors favor the expressway. One is the low cost of installing rapid transit lines in the median strip of an expressway; the other is the efficient use such lines make of the space they occupy. For instance, a double-track rapid transit line has a passenger-carrying capacity equivalent to seven 8-lane highways jammed with automobiles. The additional cost of providing a rapid transit right of way in the median strip of the Congress Street expressway in Chicago, for example, is estimated at less than 5 per cent of the total cost of the project. This method of increasing passenger-carrying capacity represents a tremendous saving in community funds as compared to the expense of building additional expressways to handle a like number of people in automobiles. Moreover, rapid transit riders do not add vehicles to further complicate the business area traffic situation.

The automobile plays an important part in this proposal; in fact, it will actually result in greater usage than if rapid

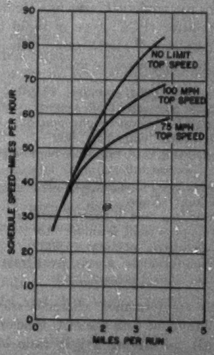


Fig. 1. Maximum theoretical performance with acceleration and deceleration rates of 3.0 mph per second

---

Paper 54-252, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 23, 1954; made available for printing May 6, 1954.

S. J. Vouch is with the General Electric Company, Erie, Pa.

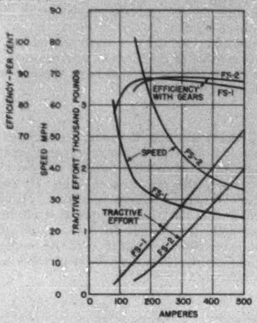


Fig. 2. GE-1247 motor characteristic curves on 300 volts, gear ratio 6.14 to 1, wheel diameter 28 inches

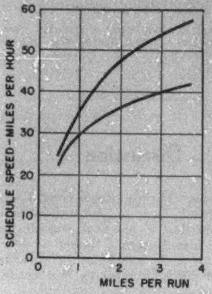


Fig. 3. Upper curve shows performance with the new GE-1247 motor, lower curve shows performance of present PCC rapid transit car

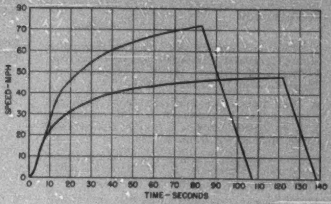


Fig. 4. Speed-time curves for 1½-mile run. Upper, GE-1247 motor. Lower, present PCC rapid transit car

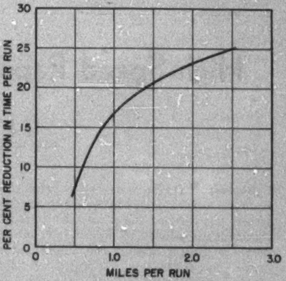


Fig. 5. Saving in time possible with new equipment compared to PCC rapid transit car

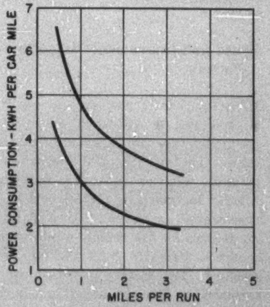


Fig. 6. Power consumption. Upper, high-speed equipment. Lower, present PCC rapid transit car

transit were not available. This follows because the automobile can be driven from outlying areas to the nearest rapid transit station by another member of the family, thus making it available for family use all day instead of its being immobilized in the business area. Parking facilities adjacent to rapid transit stations in outlying areas will provide convenient low-cost parking when desired. This co-operative use of individual and mass transit facilities uses each type to best advantage and everyone benefits because over-all time for the journey is reduced, over-all cost is lower, and fewer vehicles are operated in the congested business areas.

Performance Requirements

To determine maximum practical performance obtainable, a study was made using the modern President's Conference Committee (PCC) type of rapid transit car as a basis of comparison. The following assumptions were made:

Car weight, light, pounds: 40,000
Average passenger load, pounds: 7,000
Average loaded car weight, pounds: 47,000
Line voltage, volts d-c: 600
Accelerating rate, miles per hour (mph) per second: 3
Braking rate, mph per second: 3
Coasting: None
Leeway, per cent of time in motion: 5
Station stop time, seconds: 15
Rotary inertia, pounds: 6,000
Friction formula: Davis, 2-car train
Profile: Level tangent track

These performance characteristics are all obtainable with the PCC-type of rapid transit car having a gear ratio of 7.17 to 1 and 26-inch-diameter wheels.

After these boundaries had been established it was essential to know the best theoretical performance which could be obtained. This would indicate whether the ultimate in schedule ability was attractive enough to justify efforts to reach it. For this purpose it was assumed that an unlimited supply of power and propulsion equipment to use it would both be available. Therefore, the accelerating rate of 3.0 mph per second could be held constant as long as power was applied, and it would then be followed by deceleration at the constant rate of 3.0 mph per second. In effect, this produces a triangular speed-time curve. The maximum performance which could be obtained under such conditions is shown in Fig. 1.

It is evident from these curves that the schedule speed will not exceed 40 mph if the length of run is less than a mile. This is true even though unlimited power is available for acceleration and the train is pushed to the limit; that is, accelerating or decelerating at the maximum assumed rate whenever it is in motion. Although this represents good performance, it is definitely limited by the frequency of the stops. The use of express trains, as in New York, or skip-stop service with *A* and *B* stations, as in Chicago, results in lengthening runs to the point where the increase in schedule speed becomes still more attractive.

New Equipment Available

Since this study showed maximum performance to be worth striving for, an investigation was made to determine how closely it could be approached when actual equipment designs were used. For high performance it is extremely important to hold the car weight to a minimum. As far as propulsion equipment is concerned, it was decided to start with the traction motor frame size currently used for the PCC rapid transit car motor, and modify the design to pack all the power possible into this space.

The result was an increase in top permissible speed from 5,000 to 6,500 rpm, and an 82-per-cent increase in horsepower, with no appreciable increase in weight. Characteristics of this new

Fig. 7. Lightweight rapid transit train used in Chicago

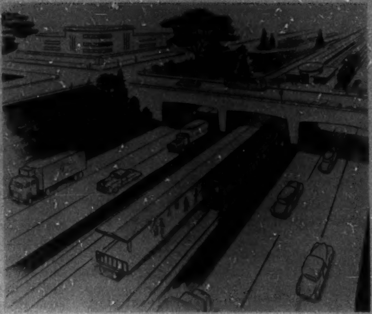


Fig. 10 (right). Artist's conception of Chicago's Congress Street Expressway with rapid transit tracks in the center mall. Parking facilities for transit riders may be provided in locations convenient to the rapid transit stations

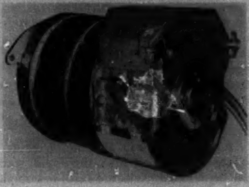


Fig. 8. GE-1247 100-horsepower 300-volt d-c high-speed rapid transit motor

motor design are shown in Fig. 2. Its performance is compared with that of the present PCC rapid transit car motor in Fig. 3. The upper curve shows the performance of a 2-car train equipped with eight of the new motors, with a gear ratio of 6.14 to 1, 28-inch-diameter wheels, and a balancing speed of 80 mph. Comparing this with the theoretical curve (100 mph top speed) shown in Fig. 1, it is found that the new design gives 83 per cent of the maximum performance obtainable under the assumed conditions in the range of 1 to 3 miles per run. This is a worth-while improvement.

The lower curve in Fig. 3 shows present PCC rapid transit car performance for the stated conditions, with gear ratio of 7.17 to 1 and 26-inch-diameter wheels. Performance of this car could theoretically be increased slightly by increasing wheel diameter to 28 inches. Where this change was made, however, it was accompanied by an appreciable increase in car weight for other reasons and it has actually resulted in a reduction of about 5 per cent in performance.

Another comparison of the difference in performance is given by the speed-time curves in Fig. 4. To reach a speed of 40 mph requires 40 seconds with the present cars, but only 15 seconds with the high-powered cars. The reduction in running time resulting from this sustained acceleration is shown in Fig. 5. Here is a powerful appeal to the prospective transit rider because he is interested primarily in getting to his destination as quickly as possible. The saving in time plus the convenience of eliminating the annoyance and expense of parking are certain to attract riders.

Power Consumption

The use of increased motive power naturally leads to the question of the power consumption required to obtain the increased performance. Fig. 6 gives a comparison between the new equipment and present equipment when performing as shown in Fig. 3. It is important to understand that the cost of operation is not proportional to the power consumption since the latter represents only 10 to 15 per cent of the annual fixed charges plus operating expense. Comparative studies have shown that the increased schedule ability of the high-speed equipment means that fewer cars are required to perform the service. As a consequence, initial investment and operators' wages are reduced. This results in a saving in over-all operating expense, the amount depending upon actual service conditions. Additional significant increases in performance or reduction in power consumption can be obtained by reducing the car weight and streamlining the car body.

Series-Parallel Control

A new control equipment, known as type *MCM*, has been developed to match the new traction motor. It employs two motor-operated controllers with cam-actuated contacts. One controller is used to set up circuits for series, series-parallel, and dynamic braking connections. The other is used to regulate resistance in the traction motor circuits for acceleration and dynamic braking. The series-parallel arrangement of circuits is desirable with high-speed equipment because it provides flexibility of operation. If operating conditions do not require top speed, it is not necessary to accelerate to full parallel, and power can thus be saved. In addition, the operator can shut off from parallel and then return to series, thus maintaining reduced car speed without having to shut off continually and reapply power.

Extended dynamic braking, supplemented by four track brakes, eliminates

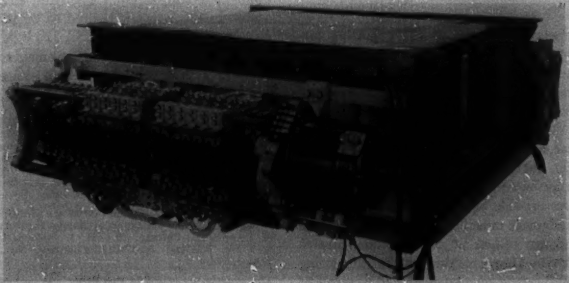


Fig. 9. Packaged control for high-speed rapid transit car

the need for air brakes. Final stopping and holding of the car are accomplished by means of electrically operated drum or disk brakes mounted on the motor shaft.

High-powered equipment such as that just described is being installed on two sample cars in Chicago, where it will be tested in anticipation of the need for high-performance cars on the center mall of the Congress Street expressway now under construction.

Conclusion

High-speed rail rapid transit can play an important part in preventing our cities from becomig enmeshed in traffic congestion to the point where they cease to function effectively. Facilities for multiple-unit trains can be provided in the center mall of expressway construction at a fraction of the cost required to build the additional highways capable of handling the movement of an equivalent number of people in automobiles. Here is a practical way to solve the traffic problem without the danger of bankruptcy in the process.

◆

Discussion

W. H. T. Holden (Consulting Engineer, Pasadena, Calif.): Particularly in the Los Angeles area the construction of rapid transit facilities is often discussed, on the basis of certain lines of thinking which are incorrect today. It is sometimes assumed that full subway under a street or the over-street elevated lines are the only alternatives. The construction proposed by the author and used on the Congress Expressway in Chicago is one example of several alternatives to the two mentioned, leading to lower capital charges for the right of way.

A further line of reasoning that has prevailed in the past and that is still heard in the Los Angeles area is based on the assumption that the user of rapid transit must walk to the station. This limits the service area of a station to approximately 1 square mile, and if enough traffic is to originate in this area to support rapid transit, very high population densities, of the order of 20,000 are required. The provision of parking facilities at stations as mentioned by the author is and has long been standard in certain main-line railroad commuter operation, as on the Main Line west of Philadelphia, and is in use on the Metropolitan Transit Authority system in Boston. This increases the service area by a factor of about fifty, provided the speed advantage of the rapid transit line is sufficient. The driver on freeways or expressways cannot keep up with a 70- to 80-mile-per-hour train paralleling the highway, while he knows that he can overtake the bus on the same roadway.

The foregoing applies only to the residential areas. At the employment areas, the automobile will not be available to the user of the rapid transit line. Station locations must be co-ordinated with office building and industrial areas, and walking plus some utilization of feeder-distributor buses must be relied upon.

The equipment described was subject to certain limitations imposed by the inadequate clearance and short radius curves of the Chicago Transit Authority rapid transit lines. The increased power consumption is not serious and is probably lower than the savings which result from higher speed. The type of acceleration discussed by Mr. Vouch in connection with Fig. 1 cannot be secured from the series motor, although it is possible with the wound-rotor induction motor.

S. J. Vouch: In these days of expensive expressway construction, any plan to increase the effectiveness of expressways in the metropolitan area should be of concern to every individual. By including provision for high-speed mass transit in the center mall, greatly increased passenger carrying capacity is obtained at relatively little increased cost. This opportunity to conserve community funds should be given serious consideration when new expressway construction is in the planning stage.

Although a certain amount of riding on surface lines which feed the rapid transit may be necessary in some cases to complete the journey, such riding would entail shorter distances than at present. It would therefore be subject to much less delay in the form of interference with other street traffic, particularly in congested areas. In the majority of cases the over-all saving in time will still be sufficient to convert automobile drivers to transit riders.

The use of wound rotor-induction motors for propulsion presents the possibility of sustaining acceleration to somewhat higher speeds but involves other complications in circuitry that make its advantages questionable. On the other hand, the series motors have a background of many years' successful operating including full service dynamic braking to standstill.

Economic Factors for Aircraft Electric Power Systems

R. M. BERGSLIEN
ASSOCIATE MEMBER AIEE

L. J. STRATTON
ASSOCIATE MEMBER AIEE

H. J. FINISON
MEMBER AIEE

AIRCRAFT electric power systems are comprised of the following four elements: prime mover, generation, distribution, and utilization. Each of these elements is interdependent and decisions affecting any one element may also influence decisions on the others. The interdependence is evident in any problems related to performance on economy of the electric power system.

Economy for an aircraft electric system cannot usually be measured in terms of money as is true for most engineering studies. Aircraft system economy must consider the effects on the entire aircraft operation. Some of the operating factors that must be used as comparison criteria in comparing several systems are: performance, reliability, vulnerability, maintenance problems, and life.

Weight is one important economic factor, and many decisions are possible using this factor. However, the other factors may be equally important but are often less capable of quantitative evaluations. Judgment must be used for many of these because of lack of adequate quantitative data. Many of the factors influencing performance can be evaluated in terms of equivalent weight. However, the factors other then performance may be equally or even more important. Unfortunately these other factors such as reliability and maintenance are often less capable of quantitative evaluation. Yet even for problems where these other less quantitative factors are important for decisions, knowledge of the relative weight economy is still helpful. The study presented in this paper is limited to weight economy.

Fig. 1 shows a simplified arrangement of a power supply system. Here a single

---

Paper 54-306, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted November 12, 1953; made available for printing May 14, 1954.

R. M. BERGSLIEN and L. J. STRATTON are with the Armour Research Foundation of Illinois Institute of Technology, Chicago, Ill.; H. J. FINISON is with the National Pneumatic Company, Inc., Boston, Mass.

The material presented in this paper was part of a study sponsored by the Power Plant Laboratory, Wright Air Development Center, Dayton, Ohio.

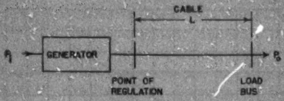


Fig. 1. Elements of a simple electric system

generator supplies power to a single load over a single transmission path. For this system, weight economy studies are necessary to determine the optimum relationships between the power supply components. For example, assuming that the safe cable temperature limits are not exceeded, the size of the cable can be decreased with a corresponding reduction in cable weight. But the effect of such a reduction is to require that more power be delivered by the generator to supply the increased cable losses. The additional power requires larger generating capacity and hence a heavier generator. It also requires use of additional fuel to generate the power and possibly additional prime mover power and weight. Hence, as the cable weight is decreased, other weight factors are tending to increase. Economic studies are necessary to establish the optimum design balance. In this paper, analysis is made of the relationships of the type just described. For this work the simplified arrangement shown in Fig. 1 and the analysis yield results of considerable interest about aircraft electric power systems. However, it must always be remembered that the actual aircraft problem may be considerably more complex:

1. Most utilization equipment does not operate continuously throughout a flight.
2. Cable losses will require an increase in generator capacity only if the losses occur at the same time that the generator is loaded to its capacity, that is, the losses will affect generator capacity only if they occur at the same time as the peak system load.
3. The load on a cable varies throughout a flight as various parts of a total load are switched on and off.
4. Dual channel service may be used for many distribution circuits to achieve higher reliability. Hence, a cable may operate at less than its rating even at normal load conditions.

Because of these and other complexities, the results obtained from the weight economy studies of simplified systems must be carefully evaluated.

Weight Data

To make a complete analysis of the weight economy of the electric power system, all of the factors which affect the weight of the system must be considered. A general discussion of some of these factors follows.

GENERATOR WEIGHT COSTS

A large number of d-c and a-c generators have been developed by industry to meet the wide requirements of aircraft applications. Thus a considerable number of data are available on the weights of specific machines. Characteristic weight curves of aircraft rotating electric machines can be determined from summarizing these data. Curves of weight per unit torque versus output torque are plotted in Fig. 2 for the following machines:

1. Blast air-cooled d-c generators, 30-volt and 120-volt.
2. Continuous duty open d-c motors, 27-volt.
3. Continuous duty totally enclosed d-c motors, 27-volt.
4. Blast air-cooled alternators, 120/208-volt 3-phase 400-cycle.
5. Continuous-duty open-induction motors, 115/200-volt 3-phase 400-cycle.

D-c generators and some alternators for aircraft applications are usually rated at constant power output over several thousand revolutions per minute. The minimum or base speed at which the machine delivers its rated output must be used in determining the output torque in order to obtain proper correlation.

It can be seen from Fig. 2 that the weight of similar machines form fairly well-defined curves. In addition, it can be seen that d-c generators and alternators fall along the same curve, namely curve A. For analytical use, it is helpful to express the weights of these source machines in equation form. The weight of d-c generators and alternators as expressed by curve A in terms of output torque is given by

$$W_g = 11.5\,T^{0.88} \tag{1}$$

The weight can be expressed in terms of power in watts, and speed in revolutions per minute as

$$W_g = 34.5(P/N)^{0.88} \tag{2}$$

If weight versus power output is plotted from equation 2 for various rated speeds, a family of weight characteristic curves results, as shown in Fig. 3.

WEIGHT COSTS OF GENERATOR CONTROL DEVICES

It is necessary to install certain controls and protective devices for each generator in order to maintain adequate control of the electric power system. The six devices usually used in conjunction with a d-c generator are: voltage regulator, field

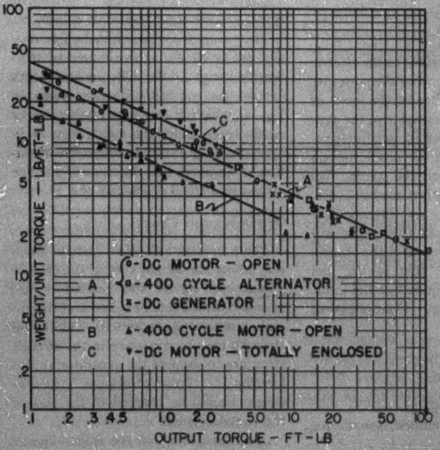


Fig. 2 (left). Torque-weight characteristics of aircraft electric machines

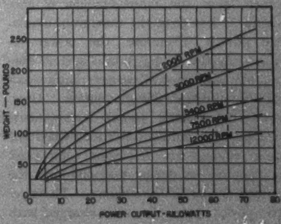


Fig. 3 (right). Speed-weight characteristic of aircraft generators and alternators

Table I. Weight of Generator Controls

System	Weight of Controls, Pounds
28-volt d-c	15.5
115-volt d-c	25.8
115/200-volt a-c	38.5

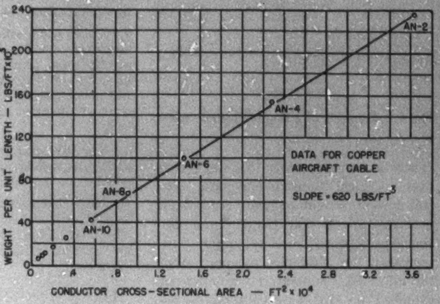


Fig. 4. Combined insulation and conductor density of copper aircraft cable

relay, differential current relay, overvoltage relay, differential voltage-reverse current relay and contactor, and circuit breaker.

Because of standardization, the combined weight of these devices is practically a fixed value for all presently used d-c generators of the same voltage. The weight of these controls represents a material part of each generator circuit, and must be charged to the cost of the electric power system.

The eight control devices usually used in conjunction with an aircraft alternator are: voltage regulator, differential current protection relay and current transformers, exciter protection relay, exciter control relay, reactive power equalizer circuit, real power equalizer circuit, circuit breaker, and frequency control circuits, that are usually a part of the drive-governor system.

Here again, because of standardization, the combined weight of these devices is practically a fixed value for all presently used aircraft 115/200-volt systems. The approximate total weight of the generator controls for the three systems discussed is shown in Table I.

Prime Mover Weight Costs

The weight costs of the prime mover chargeable to the electric power system will vary considerably from one installation to another. If the generators are driven by auxiliary power units which supply power only to the generators, the complete weight of these units and their accessories should be charged to the electric power system.

If the generators are driven directly from the main engines, the prime mover weight chargeable to the electric power system is more difficult to evaluate. During normal cruise conditions the power taken from the main engines is considerably less than the maximum available power. Therefore, the power required by the electric power system during this condition does not require any additional installed engine power. However, the electric power requirements which occur during conditions when the main engines are delivering their peak horsepower, as during take-off, limit the output thrust of the engines. The electric power system must be charged with the proportion of weight of the installed prime mover which it utilizes during the condition of peak engine horsepower.

Driving of the generators from the main engines usually requires additional installed-engine weight for gears and housing. Usually this accessory drive is in combination with drives to other engine-driven accessories. Hence, the weight of the complete accessory drive must be apportioned to the several engine-driven accessories including the generator.

If the generators are mounted remotely from the main engines and driven by an intermediate system, such as a pneumatic or hydraulic system which takes power from the main engines, the prime mover costs must include both the main engine and intermediate system charges.

Because of the numerous arrangements by which generators can be driven, it is difficult to generalize the prime mover weight costs. However, these costs are important and should be considered for each specific aircraft.

Fuel Costs

The fuel costs chargeable to the electric power system will depend on the power required by the electric loads, the losses of the electric transmission system, the generator losses, the losses in any intermediate systems between the prime movers and the generators, and the specific fuel consumption of the prime movers. Because of the variety of intermediate systems, the analysis in this paper will consider only the fuel costs of the electric power system proper, that is, the fuel costs for the power required at the input shaft of the generator. The weight of this fuel is

$$W_f = C_f t P_i \tag{3}$$

where

W_f = weight of fuel, pounds
C_f = specific fuel consumption, pounds per watt-hour
t = flight duration, hour
P_i = power input to generators, watts

Cable Weight Costs

The weight of the cables over which power is transmitted will be dependent on two factors: cable size, and type of power system. The cable weight is the combined weight of the insulation and the conductor, and may be expressed by

$$W_c = (\delta_i A_i L + \delta_c A L) \tag{4}$$

where

W_c = cable weight
δ_i = density of the insulation, pounds per cubic foot
A_i = cross-sectional area of the insulation, square feet
δ_c = density of the conductor, pounds per cubic foot
A = cross-sectional area of the conductor, square feet
L = length of the cable, feet

By grouping terms

$$W_c = (\delta_i A_i / A + \delta_c) A L = \delta A L \tag{5}$$

where δ is the combined cable density.

The value of this combined cable density can be closely approximated by plotting the weight of the cable per unit length against the conductor area as shown in Fig. 4, which is for cables with copper conductors. Since the data points lie in approximately a straight line, a linear relation is obtained, and the slope of the line gives a numerical evaluation of δ.

Other Weight Costs

There are several other factors which may affect the weight cost of the electric power system. If the generators are mounted on the main engines and the engines are wing-mounted, additional strength may be necessary in the aircraft structure to support the generator at this location. This additional structure weight should be charged to the cost of the electric system.

Space limitations in the engine nacelle

can also affect the cost of the electric system. This is particularly true when the main engines are axial-flow turbo-jet units. This type of engine has a relatively small envelope diameter. If the envelope diameter must be increased to accommodate the generator or an intermediate system which in turn drives the generator, the drag of the aircraft may be increased. This additional drag can be represented in terms of weight which should be charged to the electric system.

The method of cooling the generator is another consideration which enters into the cost of the electric power system. Most present-day aircraft generators utilize blast air-cooling. It has been shown that this method of cooling increases the drag of the aircraft, especially at the high speeds attainable with jet aircraft. The effect of this drag must be charged to the cost of the electric system. In future aircraft this method of cooling may not be utilized as extensively as it is now. New methods which may be used to cool the generators may or may not affect the drag of the aircraft. Nevertheless, the weight of the cooling system should be charged to the cost of the electric system.

Economy Study of D-C Transmission Systems

The total weight of an electric power system used to deliver power to a load is the sum of the weight of the elements of the system which has been described

$$\Sigma W = W_c + W_g + W_f + W_{pm} + W_{misc} \qquad (6)$$

where W_{pm} represents that part of the weight of the prime mover which is chargeable to the electric power system, and W_{misc} represents the other weight costs such as drag and controls.

In the simple system shown in Fig. 1, as the weight of the cable is decreased by decreasing the area, its resistance and hence its losses increase for a given power delivered and for a given length. The weight of all of the other factors in equation 6 will correspondingly tend to increase or may remain constant because of the increased power demands necessary for increased cable losses. Their weight will be dependent on the required generator power which can be expressed in terms of some system design factor such as voltage drop or current density. Voltage drop will be used in the following development.

GENERATOR WEIGHT EXPRESSED AS A FUNCTION OF VOLTAGE DROP

The power output of the generator can be expressed in terms of the power delivered to the load P_o and the power loss in the cable P_c

$$P_g = P_o + P_c \qquad (7)$$

The power loss in the cable may be expressed as a function of the voltage drop

$$P_c = (\Delta E)(I) = (\Delta E)P_o/E_o \qquad (8)$$

where ΔE is the voltage drop, and E_o is the voltage at the load. Therefore the generator output power is

$$P_g = P_o(1 + \Delta E/E_o) \qquad (9)$$

For some analyses it is desirable to have the expressions in terms of E_g rather than E_o namely

$$P_g = \frac{P_o}{1 - \Delta E/E_g} \qquad (10)$$

The weight of the generator as given in equation 2 can now be expressed in terms of the voltage drop

$$W_g = 34.5\left[\frac{P_o}{N(1 - \Delta E/E_g)}\right]^{0.86} \qquad (11)$$

FUEL COSTS EXPRESSED AS A FUNCTION OF VOLTAGE DROP

Fuel costs are determined by the power input to the generator as shown by equation 3. This can be expressed in terms of generator output power and generator efficiency, and hence in terms of the power at the load by the use of equation 10

$$W_f = \frac{C_f t P_o}{\eta_g(1 - \Delta E/E_g)} \qquad (12)$$

CABLE WEIGHT EXPRESSED AS A FUNCTION OF VOLTAGE DROP

The weight of the cable is given by equation 5 which may be manipulated so as to express the cable weight in terms of current and current density as indicated

$$W_c = \delta A L = \frac{\delta L I}{I/A} \qquad (13)$$

The voltage drop is proportional to the current density by the following relationship

$$\Delta E = \rho L I/A \qquad (14)$$

where ρ = resistivity of the cable conductor.

Using this relationship, the cable weight is

$$W_c = \delta\rho L^2 I/\Delta E \qquad (15)$$

If the current is replaced by P_o/E_o, and then E_o is expressed in terms of the generated voltage and voltage drop, the cable weight becomes

$$W_c = \frac{\delta\rho L^2 P_o}{E_g(1 - \Delta E/E_g)\Delta E} \qquad (16)$$

Economic Study of a 3-Phase A-C System

Similar weight expressions can be derived for the three-phase a-c system. However, the analysis of this system is more complex because of the power factors of the load and transmission line. To obtain convenient weight expressions for the component parts of the a-c system two assumptions must be made:

1. The power factor at the sending end of the line is equal to the power factor at the load. This assumption is quite valid for the transmission lengths involved in aircraft systems.
2. The reactance of the transmission line is small compared with its resistance. Therefore, the ratio between the resistive component of the voltage drop and the total voltage drop approaches one. This assumption is quite valid if the cables are arranged so as to obtain minimum effective spacing, as shown in Table II.

ALTERNATOR WEIGHT EXPRESSED AS A FUNCTION OF VOLTAGE DROP

The power output of the alternator in terms of the power delivered to the load P_o and the power loss in the cable P_c can be expressed as was done in equation 7 for the d-c system. The power loss in the cable may be expressed in terms of the number of phases, line current, and resistive component of the voltage drop. The resistive component of the voltage drop is assumed to be equal to the total voltage drop as mentioned in the second assumption. The expression for P_c is

$$P_c = 3(\Delta E)I = 3(\Delta E)P_o/3E_o(pf) = (\Delta E)P_o/E_o(pf) \qquad (17)$$

where

$\Delta E = \rho L\, I/A$ = resistive component of the voltage drop
I = line current, amperes
E_o = line to neutral voltage at the load, volts
pf = power factor of the load

The power output of the generator can now be expressed as

$$P_g = [1 + \Delta E/E_o(pf)]P_o \qquad (18)$$

But according to the first assumption

$$E_o(pf) = E_g(pf) - \Delta E$$

Table II. Ratio of Resistive Voltage Drop ΔE_R to Total Voltage Drop ΔE for Minimum Spacing at 400 Cycles

AN Cable Size	$\Delta E_R/\Delta E$
2	0.71
6	0.92
10	1.0
14	1.0
18	1.0

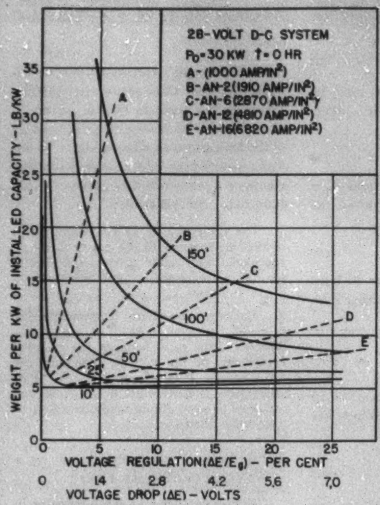


Fig. 5 (left). Weight characteristics of a 28-volt d-c system, zero flight time

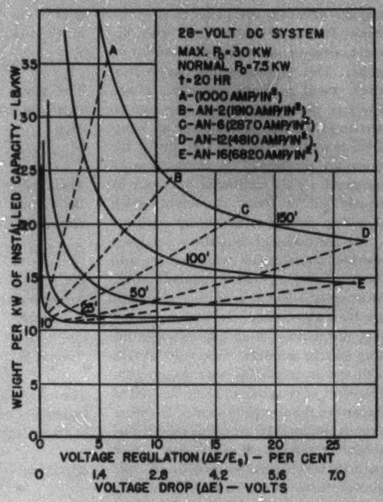


Fig. 6 (right). Weight characteristics of a 28-volt d-c system, partial load for 20 hours' flight time

and therefore the power output of the generator can be expressed in terms of E_g

$$P_g = \frac{P_0}{[1-\Delta E/E_g(pf)]} \tag{19}$$

The alternator weight can now be expressed from equation 2 as

$$W_g = 34.5\left[\frac{P_0}{N[1-\Delta E/E_g(pf)]}\right]^{0.86} \tag{20}$$

FUEL COSTS EXPRESSED AS A FUNCTION OF VOLTAGE DROP

The fuel costs can be determined from equation 13 and expressed as

$$W_f = \frac{C_f t P_0}{\eta_g[1-\Delta E/E_g(pf)]} \tag{21}$$

CABLE WEIGHT EXPRESSED AS A FUNCTION OF VOLTAGE DROP

Similarly, the cable weight for the a-c system can be expressed as

$$W_c = \frac{\delta \rho L^2 P_0}{E_g(pf)[1-\Delta E/E_g(pf)]\Delta E} \tag{22}$$

Comparison of the equations derived for both systems will show that the equations are identical except that E_g in the d-c equations are replaced by $E_g(pf)$ in the a-c equations.

Other Weight Costs

All factors except W_{misc} and W_{pm} in equation 6 have been expressed in terms of ΔE. In the analysis of the three voltage systems the only part of W_{misc} which differs is the weight of the controls.

The prime mover weight costs are made up of three components, namely

$$W_{pm} = W_L + W_{cl} + W_{int} \tag{23}$$

where

W_L = weight charges for the portion of the electric load which occurs at the same time that the main engines are delivering peak power. This weight charge is a constant for a given load

W_{cl} = weight charges for the cable losses which occur at the same time that the main engines are delivering peak power

W_{int} = weight charges of an intermediate system between the main engines and the generators

For each of the systems, the power demanded of the engines because of the electric load depends on the efficiencies of the generator and intermediate system. Although the combined efficiency of an a-c system using a constant-speed drive as an intermediate system is undoubtedly lower than that for the 28-volt and 115-volt d-c systems, for the following analysis the engine weights are assumed to be the same. However, the efficiencies are taken into account in determining the fuel charges.

Prime mover weight costs for cable losses are chargeable on only those cable losses which occur at peak engine horsepower. For many applications the cable losses at time of peak engine horsepower are small because of the light electrical load at that time. In the analysis of the next section it will be assumed that these charges are negligible.

The d-c systems in this analysis will be considered to have the generators connected directly to the main engines. Therefore, W_{int} will be equal to zero for these systems. However, the a-c system is considered to be a constant frequency system which requires an intermediate system. Thus the weight of this intermediate system must be charged to the cost of the electric power system.

Optimum Voltage Drop for Three Commonly Used Aircraft Systems

From the equations derived for the weight of the generator, cable, and fuel, the optimum voltage drop which will permit the best weight economy of a given system can be determined. However, the attainment of this optimum voltage drop for a given system may be limited by two factors:

1. Maximum allowable current density of the cable, maximum insulation temperature limit.

2. Allowable voltage regulation at the load.

To show the relationships among the optimum voltage drop, maximum allowable cable current density, and the allowable voltage regulation, a group of curves has been plotted in Figs. 5 through 10. The conditions and data from which these curves were computed are tabulated in Table III. The load conditions were selected to be representative of the actual load conditions found on military aircraft. The flight times and transmission lengths were selected to show the change in weight over a wide range of conditions.

The curves have been plotted in terms of weight per kilowatt of installed capacity versus per-cent regulation in order to have a more common denominator with which to compare systems of different voltages. The dotted curves which intersect the major curves indicate the voltage drop and system weight for a cable of a given size and transmission distance, when the cable is operated at its maximum allowable current density.

To simplify the analysis of these systems, the point of regulation is assumed to be at the terminals of the generator, that is, the analysis applies specifically to load circuits. However, it will be pointed out later that the results obtained for the load circuits also apply to the generator circuit.

ANALYSIS OF THE 28-VOLT D-C SYSTEM

The curves for the 28-volt d-c system are shown in Figs. 5 and 6. The present voltage drop requirements of this system limit the continuous voltage drop from the point of regulation to the load to 1 volt, and the intermittent voltage drop to 2 volts.[1,2] In terms of per-cent regulation these conditions are approximately 3.5 and 7.0 per cent respectively. The continuous voltage drop requirement of the 28-volt system limits the weight for the various transmission lengths at 3.5-per-cent regulation which indicates the minimum weights which can be attained without exceeding a voltage drop of 1 volt during the maximum continuous load. From Figs. 5 and 6 it can be seen that this voltage drop requirement imposes a sever-weight penalty on 28-volt d-c systems which transmit power greater than 50 feet.

The restrictions on system weight imposed by the maximum allowable current rating of bundled cable can be determined in the following manner: Consider the 50-foot curve in Fig. 6. For this curve the dotted curves of the maximum current densities of all cables of size *AN-2* and smaller lie to the right of the point where the curve intersects 3.5-per-cent regulation. This means that if the given 50-foot system is designed to meet a voltage drop of 1 volt, the maximum allowable current densities will not be exceeded even for *AN-2* cable. Therefore, for this system the allowable current ratings of these cables do not impose a restriction on the system weight. It is obvious from the curves in Fig. 6 that this same condition exists for transmission lengths greater than 50 feet.

However, consider the system with a transmission length of 25 feet, Fig. 6. For this system the dotted curves for cable sizes *AN-2* and *AN-6* lie to the left of 3.5-per-cent regulation. This means that if these cables were used, the system weight indicated at 3.5-per-cent regulation cannot be realized. If *AN-2* cable were used in this system, the weight of the system would be that indicated at the intersection of the 25-foot curve and the *AN-2* current density curve. Therefore, it is apparent that a weight saving can be achieved by using a multiple number of small cables for transmission distances less than 50 feet.[3] For longer transmission distances the voltage drop is the limiting factor and thus theoretically no weight saving can be realized by using a large number of small cables.

It is apparent in Figs. 5 and 6 that the optimum voltage drops for the longer transmission systems, 100 and 150 feet, lie off the figures to the right. Therefore, it would be impractical to attempt to attain them because of excessive regulation. Similarly the actual optimum voltage drop for 50 feet lies far to the right on the figures. However, the shape of this curve is such that the minimum system weight can very nearly be attained by using a voltage drop which is considerably less than the optimum. This voltage drop will be referred to as the practical optimum voltage drop. The optimum or practical optimum voltage drop for the smaller transmission distances are listed in Table IV for the 28-volt d-c system.

Table IV. Optimum or "Practical Optimum" Voltage Drop, 28-Volt D-C System

Transmission Length, Feet	Voltage Drop, Volts	Approximate Regulation, Per Cent
10	1.0	3.5
25	1.0	3.5
50	2.0	7.0

If the given 28-volt d-c system were designed for a voltage drop of 2 volts, it is apparent from Figs. 5 and 6 that a material saving in weight could be obtained for transmission distances greater than 25 feet.

If Figs. 5 and 6 are compared, it will be found that the shape of the curves is changed only slightly for a flight time of 20 hours at one-fourth the designed system capacity. However, the location of the curves has been shifted upward, indicating the weight of fuel consumed by the electric power system. The fact that the shape of curves is practically the same indicates that when the system is operated at only partial load for the majority of the time the effect of cable losses on the optimum voltage drop is negligible.

A considerable cable weight saving would be realized if the 28-volt d-c system considered here were designed for the normal load of 7.5 kw. However, during conditions of maximum load the voltage drop would be excessive. This condition would produce a poorly regulated voltage at the load which would impose severe requirements on the load equipment. Therefore, this does not appear to be a practical method to obtain better weight economy of the 28-volt d-c system.

ANALYSIS OF THE 115-VOLT D-C SYSTEM

A set of curves similar to those calculated for the 28-volt d-c system are shown in Figs. 7 and 8 for the 115-volt d-c system. The present voltage drop requirements of this system limit the continuous voltage drop to 4 volts and the intermittent voltage drop to 8 volts.[1] In terms of per-cent regulation these conditions are approximately 3.5 and 7 per-cent regulation respectively. The continuous voltage drop requirement again limits the

Table III. Data Used to Determine ΔE Curves

Quantity	28-Volt D-C	115-Volt D-C	115/200-Volt A-C
Maximum continuous load, P_{omax}, kilowatts	30.0	30.0	30.0
Normal continuous load, $P_{onormal}$, kilowatts	7.5	7.5	7.5
Voltage at point of regulation, E_g, volts	28.0	115.0	115.0
Base generator speed, N, revolutions per minute	3,000.0	3,000.0	5,400.0
Specific fuel consumption,* pounds per kilowatt-hour	0.95	0.95	0.95
Generator efficiency, η_g	0.85	0.85	0.85
Constant speed drive efficiency, η_g			0.74
Combined density of the cable, δ, pounds per cubic foot	620	620	620
Resistivity of copper, ρ, ohms per foot at 100° centigrade	7.47×10^{-8}	7.47×10^{-8}	7.47×10^{-8}
Power factor of the load, pf			0.75
Transmission lengths, L, feet	10; 25; 50; 100; 150	10; 50; 100; 150	10; 50; 100; 150
Flight time, t, hours	0, 20	0, 20	0, 20
Weight of controls, pounds	21.0	26.0	38.5

* Typical reciprocating engine.

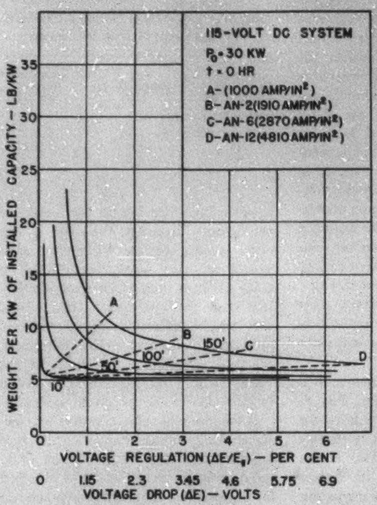


Fig. 7 (left). Weight characteristics of a 115-volt d-c system, zero flight time

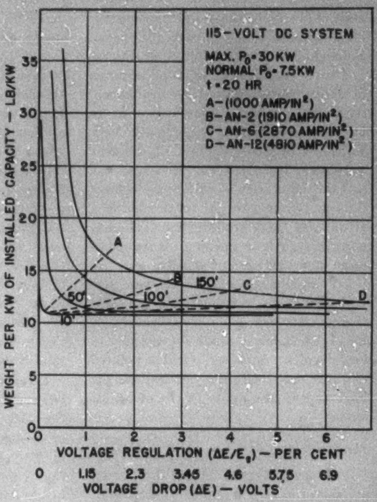


Fig. 8 (right). Weight characteristics of a 115-volt d-c system, partial load for 20 hours' flight time

minimum possible system weight at 3.5-per-cent regulation for the various transmission lengths in Figs. 7 and 8. From the figures it can be seen that this regulation imposes only a slight penalty on the 115-volt system even when long transmission distances are used.

The restrictions of the cable ratings on the system weight can be determined, as was done for the 28-volt system. However, it is apparent that for the 115-volt d-c system only a small weight saving can be made by using a multiple number of small cables as opposed to large cables. This is true for all transmission distances providing the large cables can be operated near their maximum rating.

The optimum voltage drops for the 115-volt system shown in Figs. 7 and 8 also lie off the figures in most cases. However, the curves for this system are much flatter than those for the 28-volt d-c system, thus minimum system weight can be attained at a much lower per cent regulation than was possible with the low-voltage d-c system. The practical optimum voltage drops for different transmission lengths for this system are tabulated in Table V.

Table V. Practical Optimum Voltage Drops, 115-Volt D-C System

Transmission Length, Feet	Voltage Drop, Volts	Approximate Regulation, Per Cent
10	0.5	0.4
50	3.0	2.6
100	4.0	3.5
150	6.0	5.2

From the values in Table V and the shape of the curves, it is apparent that if the 115-volt d-c system was designed for a larger voltage drop than the present specified voltage drop, no material saving in weight could be realized except where extremely long transmission distances were involved.

If the curves in Figs. 7 and 8 are compared point by point, the weight difference between curves for the same transmission length will be found to be practically a constant. This constant is the weight of fuel chargeable to the electric system. The fact that the fuel charge is nearly a constant indicates again that the cable losses have a negligible effect on the optimum voltage drop.

If the 115-volt d-c system considered here were designed for the normal system load of 7.5 kw, only a small weight saving could be realized except possibly on very large aircraft. This fact is readily apparent in Figs. 7 and 8. The flat portion of the curve for a transmission distance of 10 feet is approximately the weight of the generator in pounds per kilowatt. Therefore, the difference between this curve and any of the other curves gives the approximate weight of the cable in pounds per kilowatt. It can be seen that this difference is only a small portion of the system weight for transmission lengths up to about 100 feet. Therefore, no material weight saving could be realized for these lengths by designing for normal load. However, for transmission distances greater than 100 feet, a weight saving can be made, which though small may be desirable to attain.

ANALYSIS OF THE 115/200-VOLT A-C SYSTEM

A set of curves similar to those plotted for the d-c systems are shown in Figs. 9 and 10 for the 115/220-volt a-c system. The line to neutral voltage drop requirements of this system are the same as those specified for the 115-volt d-c system, namely, 4 volts continuous and 8 volts intermittent.[1] An analysis of these curves in the manner discussed for the other systems will yield results similar to those obtained for the 115-volt d-c system. The power factor of the load in the a-c system does not alter the general conclusions which apply to both the 115-volt d-c and 115/200-volt a-c systems. However, the effect of the power factor is evident in the wider spread of the a-c curves Figs. 9 and 10, as compared to the high-voltage d-c curves, Figs. 7 and 8. This is because of the limiting effect of the power factor on the amount of power which can be transmitted in a given a-c system. This limiting effect of power factor must be paid for in terms of weight.

It must also be remembered in comparing the conclusions of the high-voltage d-c and a-c systems that the curves for the a-c system are based on the resistive component of the voltage drop. In accepting the conclusions of 115-volt d-c system as applying to the a-c system, the

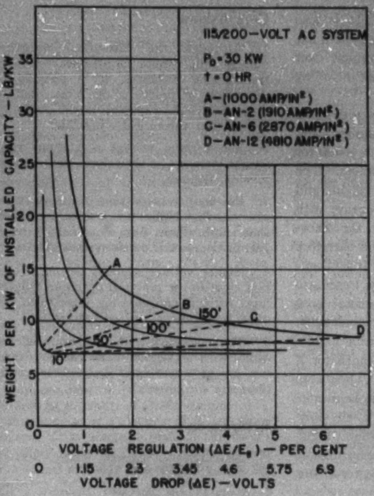


Fig. 9 (left). Weight characteristics of a 115/200-volt a-c system, zero flight time

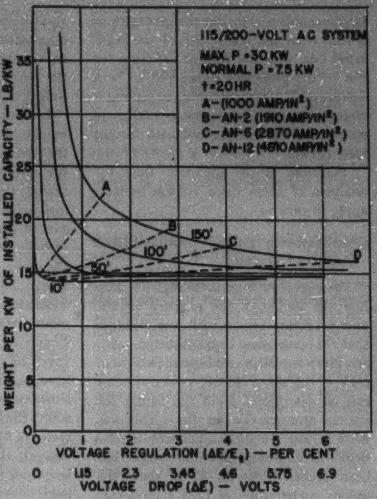


Fig. 10 (right). Weight characteristics of a 115/200-volt a-c system, partial load for 20 hours' flight time

resistive component of the transmission line impedance is assumed to be nearly equal to the total impedance. If in the majority of actual installations it is not possible to obtain this condition, it may be desirable from the viewpoint of economy to increase the allowable voltage drop of the a-c system.

General Application of the Results

Throughout this discussion, statements based on the analysis of relatively simple and specific systems have been made which, it has been inferred, apply to these systems in general. Close examination of the basic weight equations will show that these statements are not unreasonable.

Consider the expressions for the weight of the generator and fuel charges, equations 11, 12, 20, and 21. It is apparent on examination that the weight of these quantities for a given system is a function of the term $1/(1-\Delta E/E_g)$, the value of which in turn is dependent on the regulation of the system $\Delta E/E_g$. Since it is practical to operate the system only in the region where the system regulation is relatively small, the expression $1/(1-\Delta E/E_g)$ is practically equal to one in this operating range. Therefore, the generator and fuel charges are nearly constant over the practical operating range and do not affect the practical optimum voltage drop.

Now consider the expressions for the cable weight, equations 15 and 16, it is apparent from these equations that for a given system the cable weight is a function of $1/(1-\Delta E/E_g)\Delta E$. Therefore, when the regulation is small the cable weight is practically inversely proportional to the voltage drop. This indicates that the shape of the curves and thus the practical optimum voltage drop is determined by the cable in the desirable range of operation.

For different power requirements the cable weight is directly proportional to the power. This means that the shape of the curves is independent of the power requirements of the system. Therefore, the practical optimum voltage drop is constant for a given transmission length and system voltage regardless of the power transmitted.

In the simple circuit considered in the analysis, the point of regulation was taken to be at the terminals of the generator. In actual practice this rarely happens, and thus the voltage drop from the generator to the regulated bus must be considered. Since the cable is the determining factor of the optimum voltage drop, it is not

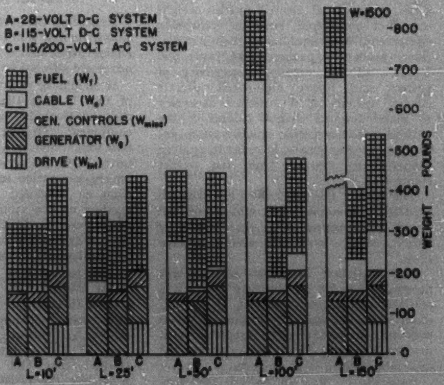


Fig. 11. Comparison of the three electric systems analyzed at 3.5 per cent regulation

unreasonable to consider the generator as a load from the standpoint of the regulated bus, with the voltage rise from the bus to the generator considered to be a voltage drop. An analysis of the generator circuit from this viewpoint would yield results similar to those obtained for the load circuits.

Therefore it is apparent that the analyses of these three voltage systems apply in general.

Comparison of Three Commonly Used Aircraft Systems

The curves of Figs. 5 through 10 show how the weight of the primary electric power system varies with the distance that the power must be transmitted. For any given regulation, the specific weight of each of the three electric power systems can be determined from the figures for various transmission lengths. This was done in Fig. 11 for 3.5-per-cent regulation. In addition, the total weight has been broken down so as to show the weight contribution of each component: the generating equipment, the cable, and the fuel.

Fig. 11 shows that for transmission distances of 100 feet, the weight of the 28-volt system, not including load equipment is almost twice that of either the 115-volt d-c system or the 115/200-volt constant-frequency a-c system. It is also apparent that the 115-volt d-c system has an appreciable weight advantage over the 115/200-volt a-c system not including weight of load equipment.

From this figure it can be seen that the weight of the generating equipment and the fuel are constant for each system for the various transmission distances. The reason that the generator weight and fuel weight are constant is because with a fixed per-cent voltage drop, the transmission losses are constant and therefore the power output of the generator is constant. Since the transmission losses are constant, the cable cross-sectional area must increase with transmission length. Equation 16 shows that the cable weight varies as the square of the length for a fixed per-cent voltage drop. Therefore, if the per-cent voltage drop is held constant, the weight of the cable is the only factor in the electric power system weight charges that varies with transmission length for a given amount of power at the load.

It is important to note that Fig. 11 shows the specific weights of the three systems only up to the load bus. For a comparison of total system weight, such items as conversion equipment, battery weight, and load equipment must be included in the weight analysis.

Conclusions

1. Although the theoretical optimum voltage drop cannot be attained for most transmission distances, it is possible to determine a practical optimum voltage drop that yields a system weight almost as low as that theoretically possible.

2. For a given voltage regulation, cable weight increases as the square of the transmission distance.

3. For long transmission distances, it is essential that higher voltages be used such that cable weight represents only a small part of the electric power system weight.

4. Higher voltage systems permit operation close to the practical optimum voltage drop without encountering limitations because of excessive voltage regulation.

5. There appears to be limited need for cables with higher allowable current densities except where high ambient temperatures are encountered.

References

1. Installation of Aircraft Wiring. *Military Specification MIL-W-5088*, March 14, 1951.

2. Standard for Aircraft Direct-Current Apparatus Voltage Ratings. *AIEE Standard No. 700*, July 1947.

3. Save Weight by Adding Cables, H. F. Rempt. *Aero Digest*, New York, N. Y., vol. 59, August 15, 1944.

4. Effects of Terminal Voltage, Load Current, and Minimum Rotor Speed on the Weight of D-C Aircraft Generators, D. H. Scott. *AIEE Transactions*, vol. 71, pt. II, Sept. 1952, pp. 191–97.

◆

Discussion

Thomas W. Armstrong (Boeing Airplane Company, Seattle, Wash.): This paper presents an interesting study of the effect of distribution system voltage drop on the generating system weight. Care must be taken, however, not to draw hasty conclusions from the data presented. As acknowledged in the paper, many variables affecting the total electric system weight were not considered, such as conversion equipment, load equipment, bus installations, protective devices, wire type, etc. Because of the extreme complexity of present and future high-speed aircraft electric systems it is believed necessary to make detailed weight studies of the various systems under consideration, including all components, in order to analyze properly the systems from the weight standpoint.

The title of the paper does not indicate accurately the principal subject of the paper which seems to be electric system weight versus distribution system voltage drop.

One important factor in cable weight determination for d-c systems which was omitted is conductor material, i.e., aluminum. Also the weight of insulation varies between manufacturers and careful selection can result in a significant reduction in total wire weight for a large aircraft.

Equation 11 does not include one very important factor in determining generator weight, namely safety factor. Military specification *MIL-E-7563* requires that, as a design objective in multigenerator installations, the generator capacity be at least twice the average continuous electric load.

Equation 9 should read

$$P_g = P_o\left(1+\frac{\Delta E}{E_o}\right)$$

Regarding the section on "Other Weight Costs," in case of pneumatic bleed-air drives, engine performance and efficiency may be adversely affected by air taken from engine.

It should be pointed out in the caption of Fig. 11 that the system weight given does not include the loads or conversion equipment. Graphs for total electric system weight for an aircraft may not necessarily be in the same relationship as those for the generation and distribution system alone as presented. In other words, depending on the number of a-c loads, it is quite possible that a 115/208-volt a-c electric system (including loads) would be lighter than a 115-volt d-c electric system.

With reference to the section, "Analysis of the 28-Volt D-C System" (third paragraph), theoretically, using several smaller wires instead of one larger one would require less copper for the same voltage drop. The smaller wires would have more exposed area and run cooler, thereby decreasing the resistance under stabilized load conditions. Whether a practicable weight saving could be achieved with the use of this consideration would be subject to study and test.

With reference to the section, "General Application of the Results" (fifth paragraph), the generator feeder voltage drop (from the generator terminal to the regulated bus) as contrasted to the load wiring, is limited only by the allowable maximum generator terminal voltage. Normally these feeders are designed on a current-carrying basis.

M. Schach and **D. H. Scott** (Naval Research Laboratory, Washington, D. C.): It is the purpose of this discussion to point out that the authors' assumption concerning current rating of cables is not generally applicable, and as a consequence, the relationships between optimum voltage, allowable voltage regulation and maximum allowable current density which are presented in Figs. 5 through 10 are valid only under restricted conditions. For clarity, consider an example. Suppose it is desired to transmit 30 kw over a distance of 50 feet. The weight per kilowatt (kw) as a function of line drop for a 28-volt system would be as shown by the authors' 50-foot curve in Fig. 6. The dotted curves which intersect this curve show the voltage drop and system weight for the indicated cable sizes carrying the maximum allowable currents given by Military Specification *MIL-*

W-5088 (reference 1 of the paper) for bundled cables. It should be noted that these ratings do not apply to bundles generally. They were the allowable currents measured for the case of three loaded cables in a conduit. It has been observed that they are valid for the case of three cables loaded in a bundle of 15 when the remaining 12 carry no current. For the authors' thermal assumption to hold, the cables would have to be separated into small groups. While this condition may be met in some cases, it is by no means general.

The general installation practice is to combine in one bundle the cables which connect the source with the loads located at a common point and those which lie along a common route. Under these conditions, there will be a range of permissible current values for each cable size, depending on the other currents which flow simultaneously in the bundle. An example of a part of the range in current capacity of aircraft cables is given in Table V of reference 1 of this discussion.

Suppose now it is desired to transmit 30 kw to a group of loads located at a common point at a distance of 50 feet from source. It is assumed that all the cables for these loads are bound together in one bundle. The allowable currents can be obtained from Fig. 11 of reference 2 of this discussion. If the cable weights subject to these thermal limits are now plotted on Fig. 6, the following results are found for this condition: 1. The intersection of all cable sizes fall within a narrow band (ΔE ranging from 0.7 to 0.9 volt). 2. The 1-volt line drop is thus met for all cable sizes (AN-18 to AN-1/0). It may also be shown that the line drop will not exceed the 3.5-per-cent limit in the 115-volt system for 30 kw and 150-foot line length when cables carry current according to our assumptions. It follows that cables with higher allowable current densities could be used to advantage. It is evident that the relations between line drop, allowable regulation, and maximum allowable current density are affected not only by total electric system kw but also by kw per bundle.

One comment should be made on the possibility of saving weight by adding wires. The authors suggest that when the line voltage drop falls below the allowable drop it would be economical to employ smaller wires in place of larger ones as suggested by Rempt (reference 3 of the paper). It is true that for the case of single cables in air, small cables dissipate more heat per unit surface area than do larger cables. However, when cables are tied together in a bundle this advantage is virtually eliminated. Moreover, a bundle composed of small cables has more layers of insulation to impede the flow of heat than a bundle composed of large cables having the same total cross-section area. It follows that, for the same temperature limit, adding wires will not save cable weight as a general result.

In summary it is emphasized that the maximum allowable currents of *MIL-W-5088* for bundled cables do not apply to bundles generally. The authors used these maximum allowable current values to derive the dotted lines of Figs. 5 to 10, and it follows that conclusions based on the position of those lines are valid only under very special thermal conditions.

REFERENCES

1. CONTINUOUS CURRENT CAPACITY OF BUNDLED CABLES FOR AIRCRAFT, M. Schach, L. D. Schroeder. *AIEE Transactions*, vol. 72, pt. II, 1953 (Jan. 1954 section), pp. 386–93.

2. CONTINUOUS CURRENT AND TEMPERATURE RISE IN BUNDLED CABLES FOR AIRCRAFT, Milton Schach, R. E. Kidwell, Jr. *AIEE Transactions*, vol. 71, pt. II, 1952 (Jan. 1953 section), pp. 376–84.

N. W. Bucci, Jr. (Westinghouse Electric Corporation, Lima, Ohio): The authors have presented an excellent analysis of economic factors of aircraft electric power systems based on weight. The purpose of this discussion is to point out the methods which are being pursued in order to save weight in new, constant-frequency a-c systems.

1. The integration of the generator and constant-speed drive into a single package, and the combination results in a smaller and lighter package.
2. The integration of oil supplies of the engine and the constant-speed drive, and the use of a single oil-cooling and oil-filtering system for both the engine and the drive results in weight savings.
3. The use of oil-cooling instead of blast-cooling. As pointed out in the paper, blast-cooling increases the drag of an aircraft. The effect of the drag is charged to the cost of the electric system. The use of oil-cooling will result in a lighter system when the factors of drag (owing to blast-cooling systems) and aircraft installation weight are considered.

These are a few of the methods which are being studied to make the constant-frequency a-c system more economical from an over-all weight standpoint. The trend appears to be to co-ordinated system design (engine, drive, and a-c generator) thus providing minimum complexity and weight. This trend appears applicable to systems other than the electric power system, particularly where cooling systems are involved.

Sensitivity Requirements of Reactive Load Division Circuits in Aircraft Electric Systems

EMILE S. SHERRARD
ASSOCIATE MEMBER AIEE

Synopsis: Reactive load unbalance among paralleled generators is analyzed to relate the largest unbalance with the sensitivity of the reactive load division circuit employed in the regulators and with the regulation characteristics of the individual generator-exciter-regulator units. The maximum permissible sensitivity of the reactive load division circuit is found, from stability considerations, to be determined by the unsaturated synchronous reactance of the paralleled generators. The design of the reactive load division circuit is discussed from a system viewpoint.

PARALLEL generator a-c aircraft electric systems consist of two or more generator-exciter-regulator units supplying a common bus with 400-cycle 115/200-volt, 3-phase power. Successful operation of such systems requires that the total system load be divided almost equally among the parallel generators. Real load division among the generators depends principally upon the speed-versus-torque characteristics of the individual generator drives and upon the sensitivity of any real load division circuits employed in the governors of these drives.[1,2] Reactive load division depends principally upon the regulated characteristics of each of the paralleled units and upon the sensitivity of the reactive load division circuits employed in the regulators of these units. The term "regulated characteristic" will be used hereafter to denote the relation between the average of the regulated line-to-neutral voltages and the magnitude of the balanced load at rated power factor of a particular generator-exciter-regulator unit for zero input signal from the reactive load division circuit to the regulator of

---

Paper 54-311, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted August 4, 1952; made available for printing April 26, 1954.

EMILE S. SHERRARD is with the Naval Research Laboratory, Washington, D. C.

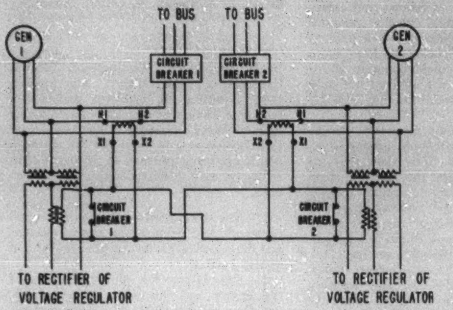


Fig. 1. Reactive load division circuit for a 2-generator system

that unit. Present specifications permit the regulated voltage of this characteristic to vary within a 5-volt band about the nominal regulated voltage. Consequently, some units of a paralleled system may show an increase in regulated voltage between no load and rated load, while others will exhibit "drooping" characteristics or "flat" characteristics over the same or a portion of the same load range. As shown in Appendix I, the worst reactive unbalance in a parallel generator system is a function of the regulated characteristics of the parallel units and the sensitivity of the reactive load division circuit. For a given set of regulated characteristics, if the worst unbalance is not to exceed a specified value, the sensitivity must not be less than a certain minimum value.

For a given system, the primary problems encountered in the application of a reactive load division circuit, or equalizer circuit, as it shall frequently be termed hereafter are:

1. The determination of the minimum equalizer sensitivity required for limiting the worst reactive load unbalance to a specified value.
2. The approximate determination of the value of sensitivity at which voltage regulator instability will occur.
3. The determination and anticipation of any undesirable effects, for example, undervoltage, that may result from faults within or malfunctioning of the equalizer circuit.

The three problems enumerated are the primary concern of this paper. No attempt is made herein to cover such subjects as: the accuracy requirements of current transformers (*CT*) employed in the equalizer circuit; the permissible phase shifts between primary current and induced secondary voltage of the equalizer transformer; or the permissible phase shift between primary and secondary voltages of the open delta transformers supplying the voltage regulator sensing circuit. Nor do the analyses in this paper consider the effect of: unbalanced generator voltages or currents; stray coupling between the reactive load division and real load division circuits; or the operation of the rectifier of the voltage regulator when the equalizer circuit signal applied to the rectifier is large in comparison with the signal applied to the rectifier by the bus voltage.

The effects just listed are peculiar to components of the system, and their magnitude may vary widely according to the manufacturer of each individual component. If these effects result in poor system performance, they must be eliminated or reduced to the point where they have only a small effect on system performance. The viewpoint of this paper is that calculations of necessary, permissible, and desirable equalizer circuit characteristics are to be made on the basis of system requirements. If presently existing components or circuits prevent the employment of these necessary and desirable characteristics, the components or circuits must be so improved that they no longer penalize system performance or operation.

As stated in the foregoing, the principal concern of this paper is the determination of equalizer characteristics in terms of system requirements. In addition, the paper contains a section describing the equalizer circuit operation and some of the design requirements for the equalizer transformer and equalizer circuit. This section is included to furnish the average reader with desirable background information. The sophisticated reader will undoubtedly be able to deduce many design details and current transformer accuracy requirements by continuing or extending the analysis of Appendix II. In the interest of homogeneity and brevity, though, such a procedure is not followed here.

Nomenclature

Subscripts $i=1, 2, \ldots, n=$ the first, second, ..., nth quantity of a collection of n quantities. $a=$an average value. $0=$a rated value

$I=$generator line current

I_d (differential current)$=$the current caused to circulate in the primary winding of an equalizer transformer by a difference between the current of its corresponding generator and the average generator line current

$L=$per unit real component of generator phase current

$I_{2i}=$secondary current of the ith current transformer of the reactive load division circuit

$R=$total impedance of the connections between *CT* secondary terminals in the equalizer loop. This impedance is assumed to be a resistance

$R_i=$real load current unbalance of the ith generator

$\rho=$per unit regulation of a generator-exciter-regulator unit when this unit operates as a single generator system. ρ is calculated as the difference between no-load regulated voltage and full-load regulated voltage when the regulator is so adjusted as to make full-load regulated voltage equal to the rated voltage of the generator

$S=$sensitivity of the reactive load division circuit. It is defined in Appendix I as the ratio of certain partial derivatives evaluated in the neighborhood of rated generator voltage and current. This sensitivity is usually measured by the slope of the droop characteristic

$U_i=$reactive unbalance current in per unit of a particular generator

$U_1=$the largest positive value of U_i in an n-generator system

$V_{in}=$no-load regulated voltage of the ith generator

$X=$per unit reactive component of generator phase current

$x_d=$unsaturated direct-axis synchronous reactance in per unit

$Z_{pi}=$impedance looking into the primary terminals of the ith equalizer transformer

Operation of Reactive Load Division Circuit

The essential circuit for accomplishing reactive load division in a 2-generator system is shown in Fig. 1. The action of this circuit decreases the excitation of a generator carrying more than average reactive current, and increases the excitation of a generator carrying less than the average reactive current of all the paralleled generators. An auxiliary circuit-breaker contact shown in the figure is open when the corresponding circuit breaker between the generator and bus is closed. Thus the circuit for a particular unit is active only when that unit supplies the bus in parallel with one or more other units. As shown in the figure, corresponding lines of each generator contain a current transformer. The secondaries of these transformers are connected in series, with additive polarities, to form a ring or loop. Across each secondary is connected the primary winding of an equalizer transformer. As shown in Appendix II, one component of current in the primary winding of a particular equalizer transformer is the so-called differential current which is proportional to the difference between the line current of its corresponding generator and the average of all the generator line currents. The voltage induced in the secondary winding by this differential current is applied, together with voltages proportional to the line-to-line voltages of the bus, to the rectifier of the voltage regulator. Reference 3 gives a good physical explanation of the action of this induced voltage in correcting reactive current unbalance among paralleled units.

For this circuit, the rectifier output is proportional to the arithmetical sum of the 3-phase voltages at the a-c input terminals of the rectifier. These voltages are the phasor sum of the 3-phase secondary voltage of the open delta transformer and the voltage induced by the differential current in the secondary winding of the equalizer transformer. Reactive and real current unbalance both vary the phasor sum. However, only reactive current unbalance varies the arithmetical sum. Consequently, the output of the rectifier varies with reactive current unbalance but is independent of real current unbalance. If a generator's reactive current is greater than the average value, rectifier output increases and the generator's excitation decreases. If a generator's reactive current is less than the average value, rectifier output decreases and generator excitation increases. The change of generator excitation in each case reduces reactive unbalance among the paralleled generators. The sensitivity of the equalizer circuit, and the worst reactive unbalance among the parallel generators is discussed in detail in Appendix I.

The equalizer circuit sensitivity increases with a decrease in the secondary to primary turns ratio of the current transformer, and increases with an increase in the mutual inductance of the equalizer transformer. This sensitivity may be evaluated from the droop characteristic of a single generator system. The droop characteristic for a given unit is determined with the equalizer loop connections open-circuited so that the total secondary current of the current transformer circulates through the primary of the equalizer transformer. Increments of reactive load are applied to the single generator and for each load, the regulated voltage measured. The reactive current in the equalizer transformer primary causes the regulated voltage to decrease as reactive current increases. The regulated voltage in per unit plotted versus the reactive current in per unit will be nearly a straight line. The magnitude of the slope of this line is equal to the sensitivity S of the reactive load division circuit (Appendix I).

The results of Appendix I may be used to calculate the necessary sensitivity for a specified maximum reactive current unbalance and for assumed regulated characteristics. The method of calculation is illustrated by the two examples which follow.

In the case of two generator-exciter-regulator units operating in parallel, the assumption is made that one unit when operating as a single generator system has a rise in regulated voltage between no load and full rated load of 2 per cent or 0.02 per unit. The regulation of the second unit is assumed to have a droop of 0.03 per unit. Also it is assumed that no-load voltages of the single generator units are adjusted so that their difference is 0.02 per unit. Calculations employing equation 8 and using $X_0=0.66$ per unit as the rated reactive current, yield $SU_1=0.05$. Here S is the sensitivity of the reactive load equalizer circuit and U_1 is the maximum reactive current unbalance in the two generator system. If U_1 is to be no greater than 0.1 per unit, S must be equal to or greater than 0.5 per unit. If $S=0.10$ to 0.15, U_1 varies between 50 and 30 per cent.

In the case of a multiple-generator system, the generator with the worst unbalance is assumed to have a rise of 0.02 per unit, and its no-load terminal voltage is assumed to be 0.01 per unit greater than the average terminal voltages of the generators. Assuming that the average regulation of the generators is a droop of 0.02 per unit and rated reactive current is 0.66 per unit, equation 9 yields $SU_1=0.075$. The sensitivity required for maximum reactive load unbalance $U_1=0.10$ is, in this case, 0.75 per unit. If S varies between 0.10 and 0.15 per unit, U_1 varies between 75 and 50 per cent.

The values of regulation employed in these calculations might be considered extreme values for carbon pile regulators. Of particular interest are the calculated unbalances when S lies between 0.1 and 0.15, since this range of S is employed in most systems in use at present.

Permissible Sensitivity of Reactive Load Division

The permissible sensitivity of the reactive load division circuit is determined by the value of sensitivity at which sustained system oscillation occurs, the cost in size and weight of the component changes necessary for obtaining high sensitivity, and the drop in regulated bus voltage that results when the equalizer loop connections between current transformer secondary terminals is open-circuited. The following analysis determines the permissible sensitivity as dictated by the stability requirement. The weight and size increases necessary in the elements of a high sensitivity system are considered negligible. The drop in regulated bus voltage caused by an open-circuited equalizer loop must be correlated with the gain in system capacity and with the prevention of sustained generator overloads. This correlation is performed best by the designer and user of each individual system. Consequently, only the permissible sensitivity as limited by stability considerations will be considered here.

Experience with voltage regulator instability in parallel generator systems reveals that instability is usually either the single-generator or the parallel-generator type. The single generator form of instability is sometimes encountered in systems employing only one generator. In such systems continuous self-sustained voltage oscillations are most likely to be encountered for high revolutions-per-minute and no-load or light-load conditions of operation. For such operating conditions, the sensitivities of the regulator, exciter, and generator are a maximum or near maximum value and the system is unstable or has the least margin of stability.

In a single-generator system, the bus voltage will be amplitude modulated in an approximately sinusoidal or an exponential fashion by the machine's field current which is itself varying in an approximately sinusoidal or exponential fashion. The field current variation is produced by the periodically varying voltage of the exciter, in whose field the regulator resistance is varying between a large or infinite value and a smaller value that may approach its minimum value. The variation in regulator resistance is itself sustained by the input to the regulator voltage coil. This input consists of the rectified bus voltage and the output of the stabilizing transformer used to feed back the exciter terminal voltage into the regulator input circuit. In a typical system all elements except the regulator be-

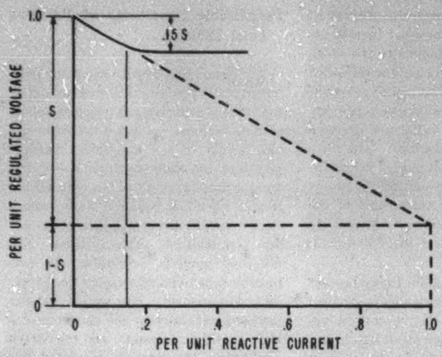


Fig. 2. Suggested droop characteristic for minimizing effect of open equalizer loop

have in a manner describable by linear differential equations.

If a generator-exciter-regulator unit exhibiting single-generator instability is paralleled with other units which are stable when isolated from or paralleled with each other, the parallel system frequently exhibits instability of the type described in the foregoing. If all the generators are stable when isolated from each other, the parallel system may be either stable or unstable. If unstable, the oscillation is usually of the parallel generator type. For such an oscillation the bus voltage and load current will be unmodulated or only slightly modulated, whereas the generator currents will exhibit a much greater degree of modulation. As before, generator field current, exciter terminal voltage, and regulator resistance will exhibit periodic variations, and the regulator might be considered the only nonlinear element of each generator-regulator-exciter unit. However, in this case the modulated generator line currents rather than the relatively unmodulated bus voltage constitute the principal portion of the varying input to the rectifier of the voltage regulator. The differential current in the primary of each equalizer transformer is equal to the difference between the line current of its corresponding generator and the average of all the generator line currents. If load current is divided equally among the paralleled generators, the differential current is equal to the totally modulated portion of the line current of its corresponding generator. In this case, only the reactive component of the totally modulated component of generator line current is effective in varying the output of the rectifier of the voltage regulator.

Whether or not parallel generator instability exists depends, in part, upon the sensitivity of the equalizer circuit. If this sensitivity is taken large enough, parallel instability occurs. If it is taken very small, parallel instability is avoided but reactive load division among the generators may be very poor. It is desirable to know a value of equalizer circuit sensitivity for which parallel generator instability will not occur in a normal system. Then a sensitivity equal to some fraction of this value may be used to obtain a compromise between accurate reactive load division and the possibility of parallel generator instability. The transfer function of a parallel system subject to parallel-generator oscillation is developed in Appendix III. A comparison is made with the transfer function of a single-generator system subject to single-generator oscillation. The single-generator system is assumed to be stable at the revolutions per minute of the parallel system and the light-load operating condition. The requirement that the transfer function describing the parallel oscillation, from a stability viewpoint, be as good as or better than the transfer function describing single generator oscillation yields the relation $S=x_d$. S is the sensitivity of the reactive load division circuit and x_d the direct axis synchronous reactance. Typical values of x_d as given in an unpublished paper by Holloway and Larson are 2.0 per unit. Thus, a value for the maximum permissible sensitivity of the reactive load division circuit is $S=2.0$ per unit. The stability margin or relative stability of the system against parallel generator instability should be as great as or greater than the stability margin or relative stability of a single-generator system.

The relation $S=x_d$ is not generally applicable unless the generator is unsaturated for light-load operation. If the generator is a constant-speed machine which exhibits some saturation at no load and rated speed, an incremental value rather than the unsaturated value of x_d should be used to determine S.

Design of Reactive Load Division Circuit

The two principal elements of the reactive load division circuit are the current transformer and the equalizer transformer. At present, it is considered desirable to use, wherever possible, the same *CT* rating throughout the system. Thus the same rating of *CT* will be used for instrumentation, differential-type protection, and reactive load division. Also, interchangeability between *CT*'s of different manufacture will be required.

A given sensitivity S of the reactive load division circuit requires that, for a given reactive current unbalance, a certain magnitude of voltage shall be induced in the secondary of the equalizer transformer by the differential current in the equalizer transformer primary. Since the rated secondary current of the *CT*'s is 2 amperes, the differential current magnitude is twice the per-unit value of generator current unbalance. Thus, if the magnitude of induced secondary voltage for a given per-unit generator current unbalance is known, the mutual reactance of the equalizer transformer is determined. Since the burden presented to the *CT* by the equalizer transformer cannot be too high, the driving point impedance looking into the equalizer transformer primary is limited in magnitude to approximately nine times the maximum permissible burden of the *CT* (see Appendix II).

To make the equalizer loop operation independent of the cyclic variation in secondary impedance caused by rectifier operation, the exciting impedance of the transformer primary should be small in comparison with the impedance across the secondary terminals when this impedance is referred to the primary.

When an equalizer loop connection between two terminals of different *CT*'s is opened, the equalizer current becomes proportional to its generator current rather than the difference between its generator current and the average generator current. This malfunctioning of the equalizer loop may result in an undesirable or dangerous reduction of the regulated bus voltage. It seems desirable to make this depression of bus voltage as small as possible by specifying the saturation properties of the equalizer circuit. In a normal system the maximum reactive or real current unbalance will be approximately 10 per cent. Therefore, the maximum differential current for normal operation will be approximately 0.28 amperes. For an open-circuited equalizer loop, the differential current will be 2.0 amperes for 1 per-unit generator phase current. Thus, it seems possible to specify that the equalizer transformer iron must saturate for primary currents greater than 0.28 amperes.

This saturation property may easily be checked by an examination of the droop characteristic. The suggested form for this characteristic is shown in Fig. 2. As shown in this figure, the regulated bus voltage is reduced linearly from its no-load value until the reactive generator current is approximately 0.15 per unit. At this value, the equalizer transformer saturates and there is little or no further reduction in bus voltage with increase of reactive load. This type of characteristic will give the minimum bus voltage reduction for an open-circuited equalizer loop.

The requirements of mutual reactance, low primary exciting impedance, large primary to secondary turns ratio (Appendix II), and saturation at 0.3 amperes primary current will probably prevent a minimum weight design of the equalizer transformer. Since present transformers weigh only a fraction of a pound, the increase in equalizer transformer weight will be more than justified by the increase in system performance and reliability.

Conclusions

The sample calculations in the body of this paper show that presently used values of sensitivity are too small for good reactive load division without frequent supervisory control by aircraft personnel. If present values of sensitivity are to be used, better regulators are required or flight personnel must make frequent adjustments for the reactive current balance.

The reduction in regulated bus voltage resulting from an open-circuited equalizer loop may be minimized by causing the equalizer transformer to saturate whenever the differential current in the primary of this transformer exceeds the maximum value obtained with normal operation of the equalizer circuit. The droop characteristic seems to be the most convenient measure of the saturation characteristics and the sensitivity of the reactive load division circuit.

From the body of the paper and the foregoing discussion, these conclusions follow:

1. In systems employing carbon pile regulators, theoretical calculations show that the employment of the presently used value of 0.10 to 0.15 per unit for equalizer sensitivity may result in reactive current unbalances of 30 to 75 per cent.
2. The maximum permissible sensitivity of the reactive load division circuit is approximately equal to the unsaturated synchronous reactance of the generator.
3. The slope of the droop characteristic and the reactive current at which saturation occurs in this characteristic should form the basis for testing the performance of the reactive load division circuit.

Appendix I. Relation Between Reactive Unbalance and Sensitivity of Equalizer Circuit

The problem of reactive unbalance among paralleled generators is complicated by the effect of the voltage regulators upon the field currents of the paralleled generators. If the regulator does not appreciably affect the generator excitations, the analysis of the problem is wholly one involving generator characteristics. If the voltage regulator is a perfect one (say, terminal voltage is maintained constant to within 0.1 per cent or better for the rated load and speed range of the generator), the analysis of the unbalance is independent of generator characteristics. Instead it is wholly dependent on the characteristics of the reactive load division circuit and the varying regulator adjustments, which cause different regulators to tend to maintain different bus voltages.

When the regulators are neither ineffective nor perfect, as is the case for aircraft type regulators, the inability of the regulators to maintain constant bus voltages, the different adjustments of the regulators, and the characteristics of the reactive load division circuit must all be considered in the analysis of reactive unbalance. A natural tool for this analysis is the feedback amplifier theory. This theory, applied to the steady-state characteristics of the paralleled generator-exciter-regulator units, is applied in the following to obtain an expression for the largest positive reactive current unbalance in terms of the imperfect behavior of the voltage regulator and the characteristics of the reactive load division circuit.

Consider any one of the n paralleled generator-exciter-regulator units. For constant speed and balanced loads, V_g, the generator terminal voltage is determined by the field excitation, by X, the reactive current, and by L, the real current. If hysteresis is ignored, the field excitation is a single valued function of V, the bus voltage input to the regulator, A, the adjustment of the regulator, and I_{dx}, the reactive component of the differential current flowing through the primary of the equalizer transformer. This functional relation is written $Vg=E(X, L, V, I_{dx}, A)$. The open-loop relation for the generator-exciter-regulator unit is given by

$$\delta Vg=\frac{\partial E}{\partial X}\delta X+\frac{\partial E}{\partial L}\delta L+\frac{\partial E}{\partial V}\delta V+\frac{\partial E}{\partial I_{dx}}\delta I_{dx}+\frac{\partial E}{\partial A}\delta A \quad (1)$$

When in parallel with other machines, δV and δVg are equal to the incremental bus voltage of the system, δVp. For an ideal equalizer circuit δI_{dx} is proportional to $\delta(X_i-Xa)$ where Xa is the average reactive current of the paralleled generators. Also

$$\frac{\partial E}{\partial I_{dx}}\delta I_{dx}=\frac{\partial E}{\partial(X_i-X_a)}\delta(X_i-X_a)$$

Therefore, equation 1 for the ith machine in the parallel system becomes

$$\delta V_p\left(1-\frac{\partial E_i}{\partial V_i}\right)=\frac{\partial E_i}{\partial X_i}\delta X_i+\frac{\partial E_i}{\partial L_i}\delta L_i+\frac{\partial E_i}{\partial(X_i-X_a)\delta(X_i-X_a)}+\frac{\partial E_i}{\partial A_i}\delta A_i \quad (2)$$

Let V_0, X_0, L_0, A_{i0}, denote the values for which the ith machine is operating at rated voltage and current with rated system load connected to the bus and with zero differential current in the equalizer transformer primary. For a small region about this operating point $\delta(X_i-X_a)=(X_i-X_a)-(X_0-X_0)=X_i-X_a$, and equation 2 may be written

$$(V_p-V_0)\left(1-\frac{\partial E_i}{\partial V_i}\right)=\frac{\partial E_i}{\partial X_i}(X_i-X_0)+\frac{\partial E_i}{\partial L_i}(L_i-L_0)+\frac{\partial E_i}{\partial A_i}(A_i-A_0)+\frac{\partial E_i}{\partial(X_i-X_a)}(X_i-X_a) \quad (3)$$

U_i, the reactive current unbalance, is defined by $U_i=(X_i-X_a)/X_a$. Equation 3 will be written in terms of U_i. The first step in this process is to approximate A_i-A_0 as follows.

The incremental bus voltage of a single-generator system operating in the neighborhood of V_0, X_0, L_0, and A_0, and with a short circuit across the primary of the equalizer transformer is evidently given by the following

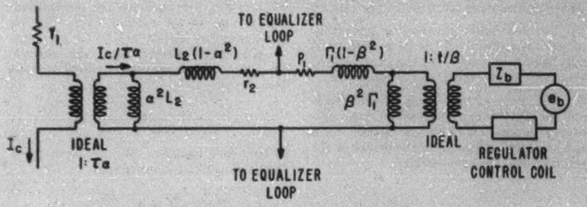
$$\delta V_{gi}\left(1-\frac{\partial E_i}{\partial V_i}\right)=\frac{\partial E_i}{\partial X_i}\delta X_i+\frac{\partial E_i}{\partial L_i}\delta L_i+\frac{\partial E_i}{\partial A_i}\delta A_i$$

If the power factor of the bus load is independent of V_g and the slight speed changes of the unit with real load, then, for per unit quantities, $\delta X_i=\delta V_{gi}\sin\theta$, $\delta L_i=\delta V_{gi}\cos\theta$, where $\cos\theta$ is the rated power factor of the generator. Therefore, in the rated load region

$$\frac{\delta V_{gi}}{\delta A_i}=\frac{\dfrac{\partial E_i}{\partial A_i}}{1-\dfrac{\partial E_i}{\partial V_i}-\dfrac{\partial E_i}{\partial X_i}\sin\theta-\dfrac{\partial E_i}{\partial L_i}\cos\theta}$$

For no load operation, $\dfrac{\delta V_{gi}}{\delta A_i}=\dfrac{\dfrac{\partial E_i}{\partial A_i}}{1-\dfrac{\partial E_i}{\partial V_i}}$

The right members of the preceding two equations are very nearly equal, although the numerators and denominators may differ greatly. This follows from the fact that 1, $\dfrac{\partial E_i}{\partial X_i}$, and $\dfrac{\partial E_i}{\partial L_i}$ are all insignificant in comparison with $\dfrac{\partial E_i}{\partial V_i}$. Also, saturation or variable sensitivity has the same effect



α AND β ARE CT AND EQUALIZER TRANSFORMER COUPLING COEFFICIENTS.
L_2, r_2, r_1 ARE RESPECTIVELY SECONDARY INDUCTANCE, SECONDARY AND PRIMARY RESISTANCES OF CT.
Γ_1 AND P_1 ARE PRIMARY INDUCTANCE AND RESISTANCE OF EQUALIZER TRANSFORMER. τ AND t ARE TURNS RATIO OF CT AND EQUALIZER TRANSFORMER.

Fig. 3. Equivalent circuit for current-transformer equalizer-transformer section of equalizer loop

on $\frac{\partial E_i}{\partial A_i}$ and $\frac{\partial E_i}{\partial V_i}$. At no load δV_{gi} is equal to $V_{in}-E_{in}$, where V_{in} is the no-load voltage of the single-generator system with regulator adjustment A_i and E_{in} is the no-load voltage of the same system with regulator adjustment A_0. Also $E_{in}=V_0+\rho_i$, where ρ_i is the regulation of the single-generator system between no load and full load. Writing a for the right members of the above equation, we have

$$a(A_i-A_0)\cong V_{in}-V_0-\rho_i \tag{4}$$

In equation 3, (X_i-X_0) may be written $(X_i-X_a)+(X_a-X_0)=U_iX_a+(X_a-X_0)$. $n(X_a-X_0)$ is the increment of reactive bus current, nX_a, in the vicinity of nX_0. nX_a is a function of bus voltage, say $f(V_p)$. Then $(X_a-X_0)=\frac{1}{n}\frac{\partial f(V_p)}{\partial V_p}\delta V_p$; and similarly $L_i-L_0=(L_i-L_a)+(L_a-L_0)=R_iL_a+\frac{1}{n}\delta g(V_p)$, where $R_i=(L_i-L_a)/L_a$ is the real current unbalance, L_a is the average real current, and $g(V_p)$ is the functional relation between L_a and V_p. Using these expressions for X_i-X_0 and L_i-L_0 and equation 4 in 3 yields

$$\delta V_p\left(1-\frac{\partial E_i}{\partial V_i}-\frac{1}{n}\frac{\partial E_i}{\partial X_i}\frac{\partial f(V_p)}{\partial V_p}-\frac{1}{n}\frac{\partial E_i}{\partial L_i}\frac{\partial g(V_p)}{\partial V_p}\right)=\frac{1}{a}(V_{in}-V_0-\rho_i)\frac{\partial E_i}{\partial A_i}+\left(\frac{\partial E_i}{\partial X_i}+\frac{\partial E_i}{\partial(X_i-X_a)}\right)U_iX_a+\frac{\partial E_i}{\partial L_i}R_iL_a \tag{5}$$

which for convenience is written

$$\delta V_p=(V_{in}-V_0-\rho_i)q+S_iU_iX_a+\rho_i'R_iL_a \tag{6}$$

where S_i denotes the equalizer circuit sensitivity, and ρ_i' and q are used to denote the ratio of various partial derivatives. Variation in ρ_i' between units is caused by generator variations affecting the numerator $\frac{\partial E_i}{\partial L_i}$ in a different manner than generator, exciter, and regulator variations affect the denominator of ρ_i'. S_i has nearly the same value for different units for the same reasons that a in the foregoing is nearly constant. Hence the subscript on the S_i's of the various units will hereafter be dropped. Because q is a little less than 1, it will hereafter be considered as equalling 1.

Equation 6 is satisfied for $i=1, 2, \ldots, n$. Also $\sum_{i=1}^{n}U_i=\sum_{i=1}^{n}R_i=0$. Let $i=1$ denote the unit with the greatest reactive unbalance. Then if the $n-1$ equations obtained from equation 6 for $i=2, 3, \ldots, n$ are added together and subtracted from $(n-1)$ times equation 6 for $i=1$, δV_p and V_0 disappear from the result which is

$$U_1S=\frac{(n-1)\rho_1-\sum_{2}^{n}\rho_i+\sum_{2}^{n}V_{in}-(n-1)V_{1n}+L_a\sum_{2}^{n}\rho_i'R_i-(n-1)L_a\rho_1'R_1}{nX_a} \tag{7}$$

In equation 7 the terms in L_a can usually be neglected since the R_i's are of variable sign, the maximum magnitude of R_i is $\cong 0.1$, L_a is about 0.75, and the magnitude of ρ_i' is probably less than about 0.1 (approximately twice the greatest magnitude of ρ_i). Also X_a by X_0 will be approximated. For a 2-generator system equation 7 becomes

$$U_1S=\frac{\rho_1-\rho_2+V_{2n}-V_{1n}}{2X_0} \tag{8}$$

For n infinite (a multiple-generator system), equation 7 becomes

$$U_1S=\frac{\rho_1-\rho_a+V_{an}-V_{1n}}{X_0} \tag{9}$$

where ρ_a is the average regulation of all the generators and V_{an} is the average no-load voltage of all the generators.

The derivatives in the expression for S should be evaluated in the neighborhood of rated load. However, no-load values for these derivatives will give a good approximation for S since these derivatives enter as a ratio in the calculation of S. In practice S is frequently evaluated by the droop characteristic test described in the body of the paper.

Appendix II. Analysis of Reactive Load Division Circuit and Current Transformer Burden

Differential Current in Terms of CT Secondary Currents

Fig. 1 shows schematically the connections of the equalizer circuit. The current transformer is assumed to have negligible hysteresis and eddy current losses. The equalizer transformer will in most cases have an air gap in the magnetic flux path. Therefore, iron losses will be neglected in comparison with reactances. Since these transformers are considered to have negligible iron losses, they may be represented by ideal transformers with their leakage reactances referred to the equalizer loop side of the two transformers.[4] An equivalent circuit for one current-transformer-equalizer-transformer section of the equalizer circuit is shown schematically in Fig. 3.

The generator line current, I_c, through the CT primary constitutes a current source. As shown in Fig. 1, the equalizer transformer secondary is connected to the voltage control coil of the regulator by means of a 3-phase bridge-type rectifier. During a cycle of line voltage, the rectifier varies the connection between voltage control coil and the secondary of the equalizer transformer. After the rectifier connects the coil and the secondary winding, the circuit may be considered linear and the bus voltage input to the circuit represented by the source voltage e_b and the source impedance z_b. When the rectifier opens the connection between transformer and voltage coil, the secondary of the equalizer circuit is considered as being open-circuited. Except for the short intervals during which the rectifier is varying connections the circuit is linear.

An exact solution of the circuit should consider the initial conditions present after the rectifier changes from a low to a high resistance or vice versa. However, in this analysis only the steady-state solution for the network is used. It is assumed that the initial conditions in the network will be so close to the steady-state values as to make the steady-state calculations of pertinent quantities a good approximation to the actual values. The following ignores the current caused by voltage sources such as e_b acting alone in the circuit, and computes only the currents resulting from current sources such as I_c. For this calculation the voltage sources are assigned the value 0, and for convenience the driving point impedance looking into the primary of the equalizer transformer is represented by Z_p.

Ordinary circuit theory yields the circuit shown schematically in Fig. 4 as the equivalent of the circuit of Fig. 3. In this Fig., I is a mesh current traversing the primary of each equalizer transformer and all the loop connections between CT secondary terminals.

Let $I_{21}, I_{22}, \ldots, I_{2n}$ denote the secondary currents of the current transformers of units 1, 2, …, n. Then, from Fig. 4 for the contour traversed by mesh current I_{2i}

$$(\gamma_{2i}+pL_{2i}+{}_{pi}Z)I_{2i}-IZ_{pi}=\frac{I_{ci}}{\tau}\alpha pL_{2i} \tag{10}$$

For the contour consisting of the secondary winding of the current transformers and the equalizer connections between *CT* secondary terminals

$$\sum_{i=1}^{n}(\gamma_{2i}+pL_{2i})I_{2i}+RI=\sum_{i=1}^{n}\frac{I_{ci}}{\tau}\alpha pL_{2i} \tag{11}$$

In equation 11, R represents the series impedance of the equalizer loop connections and is assumed to be a resistance. Summing equation 10 from $i=1$ through $i=n$ and subtracting the result from equation 11 yields

$$I=\frac{\sum_{1}^{n}Z_{pi}I_{2i}}{R+\sum_{1}^{n}Z_{pi}}=\frac{nZ_{pa}I_{2a}}{R+nZ_{pa}} \tag{12}$$

The right member of equation 12 follows from the middle member if we assume that $\sum_{1}^{n}Z_{pi}I_{2i}=nZ_{pa}I_{2a}$ where Z_{pa} is the average of the Z_{pi}'s and I_{2a} is the average secondary current of the *CT*'s. The equalizer transformer primary current I_{di} is the difference $I_{2i}-I$. Thus

$$I_{di}=I_{2i}-\frac{nZ_{pa}I_{2a}}{R+nZ_{pa}}=I_{2i}-I_{2a}+\frac{RI_{2a}}{R+nZ_{pa}} \tag{13}$$

An ideal equalizer circuit would satisfy the relation $I_{di}=I_{2i}-I_{2a}$, i.e., the primary current of the equalizer transformer is proportional to the difference between its corresponding generator current and the average of all the generator currents. Thus, the term $RI_{2a}/(R+nZ_{pa})$ in equation 13 should be regarded as a distortion caused by the resistance of the equalizer circuit connections.

Effect of Resistance of Equalizer Loop Connections on Differential Current

The distortion term $RI_{2a}/(R+nZ_{pa})$ in equation 13 will have a maximum value for rated system load. For this operating condition, excellent reactive load division is desirable. Therefore, a limitation on the magnitude of the reactive component of this distortion term is required for an average generator current of 0.75 per-unit real current and 0.66 per-unit reactive current. If $R/(R+nZ_{pa})$ is assumed to have an angle of -90 degrees, and I_{2a} an angle of a 41-degree lag with respect to the voltage from neutral to line C of the generator, the phase angle of the distortion term is $+49$ degrees. The reactive component of the distortion term will then be 0.75 $RI_{2a}/(R+nZ_{pa})$ in magnitude. The magnitude of the reactive component of $I_{2i}-I_{2a}$ for a 10-per-cent reactive current unbalance will be $0.1\times0.66\ I_{2a}=0.066\ I_{2a}$. If the distortion term is to be no greater than 10 per cent of 0.066 I_{2a}, then $|R/nZ_a|\lessapprox 0.1\times0.066/0.75=0.01$.

From this calculation, it seems advisable to make the total impedance of the equalizer loop connection less than 1 per cent of the sum of the primary impedances of the equalizer transformers. It may be necessary to choose large wire sizes for the loop connections and to make excellent terminal connections if this result is to be obtained.

CT Burden

The burden of the current transformer is defined as the ratio of the secondary terminal voltage of the transformer to the current through the secondary winding of the current transformer. The secondary voltage is evidently equal to $I_{di}Z_{pi}$. The ratio of I_{di} to I_{2i} is no greater than approximately 0.11 (10 per cent real and 10 per cent reactive unbalance). Then the burden, if only I_{di} is considered present in the equalizer transformer primary will ordinarily be less than 0.11 Z_{pi}. For perfect real and reactive load division ($I_{di}=0$), this burden will be 0.

This calculation assumes that the currents through the secondaries of the equalizer transformers caused by the e_b's in Fig. 3 are equal to one another. If this is not the case, these unbalanced currents acting alone will cause currents to flow in the *CT* secondaries. The voltage drop equal to the *CT* secondary reactance times the unbalance current should for a complete analysis be added to the $I_{di}Z_{pi}$ voltage in computing the *CT* burden. The unbalance current in the *CT* secondary will be small in comparison with the secondary current caused by rated primary current. Its neglect will introduce only small errors in the burden calculation for large generator currents.

In calculating the burden it seems advisable to neglect the effect of unbalanced currents in the secondaries of the equalizer transformers. In the first place, these currents tend to be balanced by the excellent balance among the regulator coil currents. Secondly, if the secondary turns of the equalizer transformers are fewer than the primary turns, the magnitude of this effect is decreased. Also, the unbalance is nearly random or unpredictable. The effects of this unbalance may be reduced by making the secondary turns of the equalizer transformer less than the primary turns.

Appendix III. Restricted Stability Analysis for Determining Permissible Sensitivity of Reactive Load Division Circuit

A general formulation of the stability problem in a multiple-generator system results in an almost unmanageable set of nonlinear differential equations. A solution of the most general formulation would require the services of a computer and the solution would be given in numerical rather than symbolic form. The following analysis determines in symbolic rather than numerical form a permissible sensitivity of the reactive load division circuit. To do so, the general problem is simplified by the assumptions listed below:

1. The generators are narrow-speed range machines per *MIL-G-6099* specification and operate in parallel at 400 cycles rather than at their lowest frequency rating of 320 cycles. Speed is constant and the bus load is zero.
2. The steady-state and transient characteristics of each of the paralleled generator-exciter-regulator units are identical.
3. Generator and exciter iron are unsaturated and the equations of reference 5 employing unsaturated parameters may be used to describe the generator. In these equations stator resistance may be neglected.
4. Each of the paralleled units is stable when operated in a single generator system at 400 cycles.
5. Any instability in the system is of the parallel-generator type.
6. A frequency-response method of analysis such as used in reference 6 is adequate and the frequencies of oscillation are low enough to permit neglect of amortisseur effects.

A description of parallel-generator oscillation is given in the body of this paper. With the assumptions in the foregoing this type of oscillation requires that the balanced bus voltages be unmodulated, and that the differential current in the primary of each equalizer transformer be proportional to the totally modulated component of I_c, the phase current of its corresponding generator. Since the bus voltages are balanced and unmodulated, the zero-sequence axis voltage and current are zero; and the direct axis voltage e_d and quadrature axis voltage e_q are d-c or constant quantities. The periodic field voltage e_{fd} may be resolved into a d-c or average value and an alternating or varying component. An examination of the linear machine equations shows that all rotor

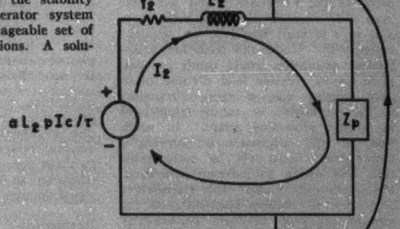


Fig. 4. Equivalent circuit for circuit of Fig. 3

currents, the direct axis current i_d and the quadrature axis current i_q will have a d-c component given by the solution of the machine equations for e_{fd}, e_d, and e_q equal to their d-c values. The varying component of these currents are given by the solution of the machine equations for e_{fd}, e_d, and e_q equal to their varying or alternating components.

The following will principally concern the varying components of machine voltages and currents, and a prime will be used to distinguish a varying component or a totally modulated component from the total value of a quantity. Thus e_{fd}' and I_c' will denote the varying component of e_{fd} and the totally modulated component of I_c. The material which follows determines the relation between i_d' and e_{fd}' for the case of parallel instability, and then determines the relation between e_q' and e_{fd}' for the case of single-generator instability.

For parallel instability the only varying voltage not equal to zero is e_{fd}'. If stator winding resistance is neglected, all the varying quadrature axis currents are zero, and the varying components of the direct axis and quadrature axis flux linkages are zero. The solution in operational form of the machine equations when the amortisseur is neglected yields

$$i_d' = X_{afd}e_{fd}' \Big/ x_d\left[R_{fd}+p\left(X_{ffd}-\frac{X_{afd}^2}{x_d}\right)\right] \tag{14}$$

I_c' is equal to $+i_d'$ cos (ωt+120 degrees). The differential current of the unit's equalizer transformer is proportional to I_c'. If the slight variation in response of the equalizer transformer to the side-band current components of the differential current is ignored, the differential current induces a totally modulated voltage in the secondary winding of the equalizer transformer. The rectifier demodulates this voltage to give an output proportional to the modulation. This output serves as an input signal to the regulator and we shall denote it by M_1. The exciter output is e_{fd}'. For a given operating point of the regulator there is a definite relation between M_1 and e_{fd}'. This relation is written

$$e_{fd}' = G_2M_1 \tag{15}$$

where G_2 will, in general, depend on both the frequencies and amplitudes of M_1. The relation between M_1 and i_d', the modulation of I_c' is given by $M_1 = SG_1i_d'$. In this expression S is the sensitivity of the equalizer circuit and G_1 is the relation between the varying component of rectifier output and any modulation of the 3-phase bus voltage about its regulated value. The over-all transfer function of the system which we denote by G^p is the product of the transfer functions in the foregoing. It is given by

$$G^p = G_2SG_1X_{afd}/x_d \times \left[R_{fd}+p\left(X_{ffd}-\frac{X_{afd}^2}{x_d}\right)\right] \tag{16}$$

The superscript p is used to denote the transfer function of a unit subject to parallel-generator oscillation.

The single-generator operation at no load will have an output voltage modulated by the varying field voltage. If amortisseur effects are neglected

$$e_q' = X_{afd}I_{fd}' = X_{afd}e_{fd}'/(R_{fd}+pX_{ffd})$$

The totally modulated components of the phase voltages will have amplitudes equal to e_q'. The rectifier demodulates the bus voltage to give a varying output proportional to e_q'. The relation between varying input to the regulator and the exciter's varying output is the same for the single-generator case as for the parallel-generator case. The over-all transfer function for the single generator case is denoted by G^s. It is given by

$$G^s = G_2G_1X_{afd}/(R_{fd}+pX_{ffd}) \tag{17}$$

Our principal concern in the following is to choose S and x_d so as to make G^p the transfer function of a stable system. G^p may be written

$$G^p = G^sS(R_{fd}+pX_{ffd}) \Big/ x_d\left[R_{fd}+p\left(X_{ffd}-\frac{X_{afd}^2}{x_d}\right)\right] \tag{18}$$

The factor $S(R_{fd}+pX_{ffd}) \Big/ x_d\left[R_{fd}+p\left(X_{ffd}-\frac{X_{afd}^2}{x_d}\right)\right]$ differentiates G^p from G^s. If S/x_d is taken equal to unity, this factor causes G^p to represent a more stable transfer function than does G^s. By assumption 4, G^s is the transfer function of a stable system. Therefore, the simple criterion $S = x_d$ assures that there is no parallel instability in the system.

References

1. Parallel Operation of Main-Engine-Driven 400-Cycle Aircraft Generators, L. G. Levoy, Jr. *AIEE Transactions*, vol. 64, Dec. 1945, pp. 811–16.

2. Governor for Variable-Ratio Transmission Used in Developmental 400-Cycle Electric System for Large Aircraft, P. F. Desch, D. E. Garr, J. G. Hutton. *AIEE Transactions*, vol. 65, Apr. 1946, pp. 194–98.

3. Alternating-Current Electrical Systems for Aircraft. Westinghouse Electric Corporation, Lima, Ohio; pp. 4–3, 4–4.

4. Communication Networks (book), E. A. Guillemin. John Wiley and Sons, Inc., New York, N. Y., vol. 2, 1947, pp. 151–156.

5. Synchronous Machines (book), C. Concordia. John Wiley and Sons, Inc., New York, N. Y., 1951, equations 32 through 40.

6. A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms, Ralph J. Kochenburger. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

Discussion

Russell W. Stineman (Boeing Airplane Company, Seattle, Washington): The author's discussion of a saturating equalizer transformer did not fully evaluate a number of objections to such a scheme:

1. Because of the physical arrangement of the components and wire size used in an equalizer loop, the most probable type of fault would be an open-circuited loop. Under this condition, transformer saturation would result in complete loss of reactive load division, probably causing one or more alternators to be overloaded or to be tripped off by exciter protection relays. In addition, if the alternator is carrying a heavy load, there will be considerable core loss owing to extreme saturation. If overheating causes the transformer secondary to become open-circuited, the voltage regulator becomes permanently disabled and the alternator is tripped off by its overvoltage protection. It is thus possible for a single malfunction (a broken loop) to cause a complete electrical outage, or even complete failure of the electric system.

2. The core loss in saturated or nearly saturated transformers causes a phase shift between primary current and secondary voltage which increases the voltage regulators' sensitivity to real load division.

3. An unbalance of real current will saturate the transformer just as effectively as reactive current unbalance. Thus, a malfunction of the real load division means will adversely affect reactive load division.

Another means has been described of preventing excessive voltage variations due to equalizer action.[1] This scheme employs relays to isolate the alternators (without removing power from any loads) whenever an abnormally strong equalizer signal appears.

Reference

1. Differential Reactive Current Protection Relay, Russell W. Stineman. *AIEE Transactions*, vol. 72, pt. II, 1953 (Jan. 1954 section), pp. 407–12.

P. F. Boggess (Westinghouse Electric Corporation, Lima, Ohio): System requirements for reactive load division and methods for evaluating these requirements can vary. The sensitivity requirements of reactive-load-division circuits have received considerable attention. Other papers have discussed power division in parallel electric power systems and how failure of power division circuits affect performance.

The author has recommended in Fig. 2 two factors that might be specified as requirements for reactive-load-division circuits: The reactive-load-division loop should have a sensitivity of *S;* and it should saturate at 0.15 *S*.

The reactive-load-division loop sensitivity is specified in volts per ampere or per unit. The sensitivity factor as shown in Fig. 2 does not appear to be specified correctly. Sensitivity should not be specified as voltage, but as the slope of the droop characteristic. The symbol *S* has been defined correctly in the Nomenclature.

The reactive-load-division loop sensitivity as specified by Fig. 2 would limit voltage droop on an open equalizer fault.

Only a part of the requirement has been specified. Saturation in the reactive-load-division loop will, in general, cause apparent phase shift in the sensing circuit, causing the loop to be sensitive to real power. A curve of voltage droop or rise versus unity-power-factor load current should be added to Fig. 2, as this characteristic may affect reactive-power division, system voltage regulation, and system stability. As saturation usually causes increased voltage sensitivity to real-power unbalance, it might be better to design the equalizer loop so that probability of failure does not require design of saturated equalizer loops.

The reactive-current unbalance, U_1, is defined in Appendix I as

$$U_1=\frac{X_1-X_a}{X_a}$$

and for approximate results

$$U_1=\frac{X_1-X_0}{X_0}$$

The reactive-current unbalance is therefore defined as the per unit of the reactive-current rating of one machine, rather than the generator current rating. This leads to some confusion. U_1 would appear to equal 0.066 per unit of the generator current rating. If U_1 were assumed to be 0.1 per unit based on current rating of the generator, then S is approximately 0.8 per unit.

Reactive-current unbalance can be computed by another development.[1] The method used, while less general, appears easier to understand. If voltage is assumed to be controlled within ±3.5 per cent (±0.035 per unit) and the maximum reactive-current unbalance shall not exceed 10 per cent (0.1 per unit) of the a-c machine rating, the sensitivity factor required may be computed for a 2-a-c generator system by

$$0=|V_{01}|-|V_{02}|-K(|I_1|\sin\theta_1-|I_2|\sin\theta_2)$$

where

$|V_{01}|$ = reference voltage of a-c generator no. 1
$|V_{02}|$ = reference voltage for no. 2 generators
K = a constant of proportionality which is a measure of the over-all effect of the equalizer
$|I_1|$ = line current for a-c generator no. 1
$|I_2|$ = line current for a-c generator no. 2
θ = angle by which the current lags terminal voltage

From the assumptions

$$(|I_1|\sin\theta_1-|I_2|\sin\theta_2)=0.1 \text{ per unit}$$

$$|V_{01}|-|V_{02}|=0.07 \text{ per unit}$$

The sensitivity factor K is then equal to 0.7 per unit.

With development of improved voltage regulators, there has been some acceptance of the following requirement for reactive-load division. Within voltage-regulation limits of ±2.5 per cent, the maximum difference in reactive current between a-c machine in a 2-generator parallel a-c system shall not exceed 20 per cent of a single a-c generator name-plate rating. This means, if one machine is 2.5 per cent high in voltage and the other is 2.5 per cent low in voltage before paralleling, that after paralleling, one machine will supply less than 10 per cent more than its share of reactive current and the other machine will supply less than 10 per cent less than its share of reactive current. Very little is gained in defining a-c generator rating or system rating if greater sensitivity is used while the tendency of instability increases. The sensitivity factor S for the above requirement would be approximately 0.333 per unit. This value is approximately one-half the value indicated by the author's paper.

REFERENCE

1. POWER EQUALIZER SYSTEMS FOR AIRCRAFT ALTERNATORS, John A. Granath, Albert K. Hawkes. *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 209–17.

Emile S. Sherrard: I should like to acknowledge the use in Appendix III of an idea of Charles Concordia; to call attention to test data which substantiates the theoretical analyses of Appendix I and III; to suggest a more useful definition for reactive load unbalance than is given in Appendix I; and to reply to some parts of the discussions of Mr. Stineman and Mr. Boggess.

The stability analysis of Appendix III assumes that, for practical purposes, the stability of a single generator supplying an infinite bus is identical with the stability of the actual system. This equivalent circuit for judging voltage regulator stability as a function of reactive load division sensitivity is, to the best of my knowledge, the invention of Charles Concordia. Approximately 10 years ago, he employed this equivalent circuit in unpublished work that analyzed parallel generator instability in a d-c generator system employing finger-type regulators. I regret that haste in the preparation of this paper two years ago caused me to omit the proper reference to Mr. Concordia's work in the body of this paper.

Both Mr. Boggess and Mr. Stineman indicate that a saturation or limiting of the reactive load division signal may be impractical with the present circuit. If test data should confirm this possibility, I would favor the use of a reactive-load-division circuit that permits reliable limiting of the equalizer signal. In particular, I see no reason why the present circuit cannot be replaced by a circuit similar to that used for obtaining real load division in aircraft system.[1] Such a circuit change would:

1. Permit the parallel operation of generators of different manufacture. At present such operation is blocked by the interconnection of the reactive unbalance sensing circuit and the voltage sensing circuit, and the fact that different manufacturers employ different voltage sensing circuits.
2. Eliminate the jump in regulated bus voltage which occurs whenever the generator is switched from single generator operation to multiple generator operation. This jump I have found to be 1 to 2 per cent of the regulated voltage. It is caused by the change in impedance offered to the voltage sensing circuit of the regulator by the equalizer transformer whenever the primary of the equalizer transformer is connected to the equalizer loop after having been short-circuited for single-generator operation.
3. Generate a reactive-load-division signal which is independent or nearly independent of real load unbalance.
4. Permit the use of simple and reliable limiting devices such as rectifiers, and permit parallel operation of generators of different manufacture to be obtained by specifying: 1. the resistance of an equalizer winding on a carbon pile regulator or a magnetic amplifier; 2. the slope of the output of this circuit expressed as per-unit output volts divided by per-unit reactive generator current; and 3. the per-unit output voltage at which saturation occurs.

Mr. Stineman indicates an open-circuited equalizer loop might cause loss of the electrical system. I believe he will agree with me when I say that the design and application of current transformers, equalizer transformers, and voltage regulators should be such that an open-circuited equalizer loop will not cause insulation or thermal failure of any of these components. If this is the case, it seems to me that the dangerous result of an open-circuited equalizer loop is the reduction of bus voltage that occurs. To reduce the danger of depressed bus voltage, I have suggested that the equalizer signal be limited to approximately 1.5 times its maximum magnitude during normal operation. Even at the expense of a change in the present circuit, I still believe this signal should be limited.

If Mr. Boggess will examine Fig. 2, he will see that the value of S indicated is numerically equal to the sensitivity since it is the ratio of an increment of bus voltage drop to a unit increment of reactive current increase. The reader may easily verify that the definition of U_1 in Mr. Boggess's discussion is in per unit, because a division of both numerator and denominator currents by base stator current converts all quantities, including U_1, to per unit values.

In the last part of Appendix I, the sentence "Also X_a will be approximated," by X_0 may cause other readers than Mr. Boggess to approximate X_a by X_0 in the definition of U_1. I have used this approximation only in approximating X_a in the denominator of the third from last equation of Appendix I. To avoid confusion, I should like to offer a different definition of reactive unbalance which removes the need for approximation and also results in a simpler formula for reactive unbalance.

If we redefine reactive unbalance U_1' by the relation $U_1'=U_1X_a=(X_1-X_a)$, the equation for reactive unbalance becomes

$$U_1'S=(X_1-X_a)S=\rho_1-\rho_{av}+V_{an}-V_{1n} \quad (19)$$

For a 2-generator system, equation 19 has the simple form

$$2(X_1-X_a)S=\rho_1-\rho_2+V_{2n}-V_{1n} \quad (20)$$

Equations 19 and 20 permit ready calculation of reactive unbalance or sensitivity if the no-load regulated voltages and the regulation (drop in regulated voltage between no load and full load) of each of the generators are known. However, I am certain that different engineers will use these equations to calculate widely differing values for the sensitivity required for acceptable reactive load division. For example, an engineer in the Department of Defense might employ the largest value of $\rho_2 - \rho_1$ permitted by military specifications. This would be approximately 0.05 per unit if one generator had a rising regulated voltage characteristic with load and the other generator had a drooping regulated characteristic, and if the regulated voltage of each generator were required to be within ±2.5 per cent of the nominal value. Such an engineer might also assume that the 2-generator system was installed in a small or medium-sized plane and that the no-load voltages were adjusted by a radio operator or navigator to within one scale division of the indication of nominal voltage upon a 2-per-cent accuracy aircraft instrument. If the adjustments were made at different times using different voltmeters or at differing points in the warm-up time of each regulator, such an engineer might estimate that $V_{1n} - V_{2n}$ was no less than 0.05 per unit. If he desired 90 per cent of total generator reactive current to be available at the bus without exceeding the rated current of a generator, he would assume $X_a = 0.9X_0$ and $X_1 = X_0$, where X_0 is rated reactive current in per unit. These assumptions yield the relation $2S\ 0.1X_0 = -0.05\ -0.05$, and for $X_0 = 0.66$, the required sensitivity is -0.75 per unit.

On the other hand, an engineer of an equipment manufacturer might assume that the regulation of his units would differ by no more than 10 per cent of the permissible regulation of 0.025. Thus, he could assume that $\rho_2 - \rho_1$ was $0.2 \times 0.025 = 0.005$. Similarly, he might assume that the no-load voltage of his regulators was determined by a factory setting and was accurate to within ±1/2 per cent of nominal voltage. Thus, he might assume a maximum value of $V_{1n} - V_{2n}$ to be 1 per cent or 0.01 per unit. For these assumptions, the value of S necessary to insure 90-per-cent availability at the bus of installed generator reactive capacity could be calculated from $2S0.1X_0 = -0.015$ as -0.11 per unit.

If an air frame engineer were to apply this formula, he might wish to have his calculations recognize the fact that his system load is 0.95 power factor rather than the 0.75 power factor rating of his generator equipment. He might also choose a different criterion from the 90 per cent of installed capacity available at the bus criterion used in the foregoing.

An experimental check of equation 25 has been made for a 2-generator system employing magnetic amplifier regulators. The test results may be found in reference 2. These test results are in good agreement with predicted or theoretical results, except in the region of small reactive unbalances where such effects as current transformer accuracy, imperfect operation of the reactive load division circuit and phase voltage unbalance of the generators seem to affect reactive unbalance. The stability analysis in this paper has been checked experimentally by tests whose results may be found in reference 2 of this closure. They apply to a system employing carbon pile regulators and now obsolete generators.

REFERENCES

1. NEW DEVELOPMENTS IN DIFFERENTIAL-TYPE HYDRAULIC TRANSMISSIONS AND CONTROLS, **R. H. Eisengrein.** *AIEE Transactions*, vol. 72, pt. II, 1953 (Jan. 1954 section), pp. 423–29.

2. SENSITIVITY REQUIREMENTS OF REACTIVE LOAD DIVISION CIRCUITS FOR PARALLEL GENERATOR SYSTEMS, **E. S. Sherrard.** *Naval Research Laboratory Memorandum Report No. 185*, July 30, 1953.

A Self-Excited Induction Generator with Regulated Voltage

H. M. McCONNELL
ASSOCIATE MEMBER AIEE

THE synchronous generator with direct-driven exciter has several disadvantages when applied in airborne equipment. However, no competing system has succeeded in achieving a smaller weight in pounds per kva at a given speed. The practical difficulties are overcome at present by careful engineering, so that a reliable generating system is obtained. These difficulties relate to brush performance at high altitude, commutation at high speed, overhung moment due to the action of exciter, bracing of the rotor windings against centrifugal force, and high cost of manufacture.

Many of these disadvantages can be avoided if the induction generator can be adapted to automatic voltage control. There are no brushes or commutator, no overhung exciter, and the rotor structure which is so difficult to manufacture in the synchronous generator is replaced by the robust squirrel cage. This paper describes work directed toward the utilization of the induction generator in airborne service. Its field of usefulness in competition with the standard synchronous generator is discussed.

Voltage Control

When induction generators are used in central station service, excitation is supplied by synchronous generators operating in parallel at other locations on the system. Excitation for isolated service must be supplied by capacitance. Excitation control, and consequently voltage control, is obtained by varying the net capacitive reactance.

In addition to the achievement of a generator without brushes, it is desirable to eliminate both electronic tubes and moving parts from the regulating system. The control scheme to be described accomplishes both these objectives. Fixed static capacitors, connected in parallel with the generator terminals or in a series-parallel combination, provide the excitation. Self-saturating magnetic amplifiers connected across the machine terminals provide excitation control, by drawing a controllable lagging reactive current. A pilot magnetic amplifier stage controls the main magnetic amplifiers. The equipment is arranged as shown in the block diagram of Fig. 1.

The excitation system is designed to provide sufficient capacitive current drain for the most extreme condition of load, when the main magnetic amplifier is driven to minimum current. Other conditions requiring less capacitive current at the regulated voltage are obtained by increasing the main magnetic amplifier current drain. Since an increase in main magnetic amplifier current always drives the voltage down, single-valued control characteristics with respect to voltage are an inherent property of the system.

Initial voltage build-up depends on the existence of residual magnetism in the rotor iron. It has been found experimentally that the grade of sheet steel used in industrial machines will retain sufficient induction to assure build-up at no load if the inactive period has been short. However, an inactive period of a few days has been found to reduce the residual below the necessary amount. It is suggested that a few per cent of the rotor

---

Paper 54-317, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 2, 1954; made available for printing May 7, 1954.

H. M. McCONNELL is with the Carnegie Institute of Technology, Pittsburgh, Pa.

The work reported in this paper was supported by the Office of Naval Research, under contracts *N7ONR 30306* and *30308*, projects 075-272 and 275.

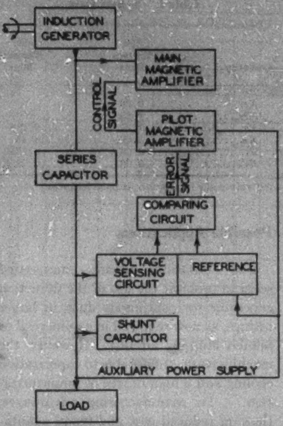

INDUCTION GENERATOR
MAIN MAGNETIC AMPLIFIER
CONTROL SIGNAL
PILOT MAGNETIC AMPLIFIER
SERIES CAPACITOR
ERROR SIGNAL
COMPARING CIRCUIT
VOLTAGE SENSING CIRCUIT
REFERENCE
SHUNT CAPACITOR
AUXILIARY POWER SUPPLY
LOAD

Fig. 1. Block diagram of self-excited induction generator with magnetic amplifier control

punchings should be replaced by a steel having properties approaching that of a permanent magnet material. Moderate success has been obtained experimentally by shortening the squirrel cage and adding an Alnico ring driven from the shaft by friction, although this cast material has definite mechanical stress limitations.

Factors of Design

There are several advantages of the induction generator, which taken together may be decisive when the limit of miniaturization is contemplated. First, the rotor ampere-wire loading is less than in the synchronous generator, while the stator ampere-wire loading is greater. Since the rotor is the most difficult to cool, this altered ampere-wire loading more nearly matches the inherent heat dissipation ability of the structure. A second factor is that the electrical insulation between rotor copper and rotor iron is an absolute minimum, the oxide film being sufficient. Consequently, heat transfer from rotor copper to rotor iron may be greatly increased. A third important property of the rotor structure is that mechanical bracing of field coils is not required. Higher peripheral speeds are possible and a saving in manufacturing cost is effected. Fourth, the squirrel cage is the same as a damper winding for operation on unbalanced loads.

These factors of design all indicate that some reduction in the weight of the generator itself may be achieved. In addition, the exciter of the synchronous generator is eliminated altogether. Balancing the weight advantage gained in this manner is the fact that the excitation and control must be provided by external apparatus, which altogether quite likely offsets the previous gain. Definite answers to this question must await the design of a system using the same materials and allowable temperature rise as are employed at present in synchronous generators.

Factors of Operation

The most serious limitation of the induction generator system is its inability to deliver sustained short-circuit current. The present philosophy of circuit protection in aircraft requires that faults be cleared by limiters. Consequently the generator system must deliver a specified short-circuit current for a specified time. The induction generator with capacitor excitation behaves on short circuit somewhat like the d-c shunt generator, with the exception that the induction generator time constant is so short that no appreciable heating of a fuse element may be obtained from its transient current.

A similar remark applies to the capability of carrying momentary low-power factor overloads, such as starting an induction motor. This kind of load acts to remove excitation just as does the main magnetic amplifier. If the control circuit is not fast enough to transfer the reactive current from main magnetic amplifier to

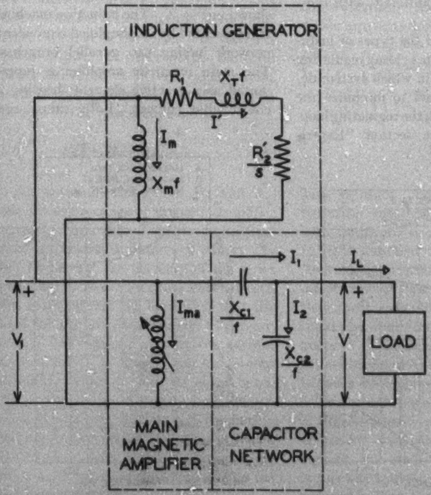

INDUCTION GENERATOR
R_1
$X_T f$
I'
I_m
$X_m f$
$\frac{R_2'}{s}$
I_1
I_L
+
V_1
I_{ma}
$\frac{X_{C1}}{f}$
I_2
$\frac{X_{C2}}{f}$
+
V
LOAD
MAIN MAGNETIC AMPLIFIER
CAPACITOR NETWORK

Fig. 2. Equivalent circuit of generator with capacitors and main magnetic amplifier

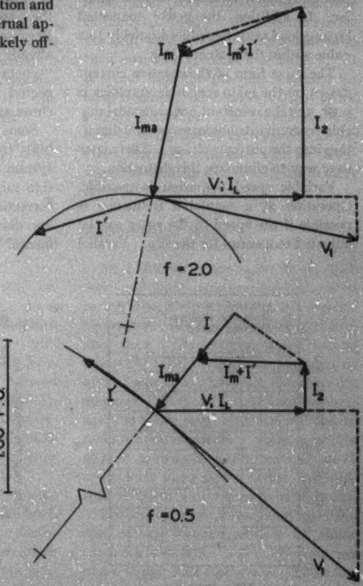

I_m
I_m+I'
I_{ma}
I_2
V; I_L
I'
V_1
f = 2.0
1.00 P.U.
I
I_{ma}
I_m+I'
I'
I_2
V; I_L
f = 0.5
V_1

Fig. 3. Phasor diagrams, unity power factor load

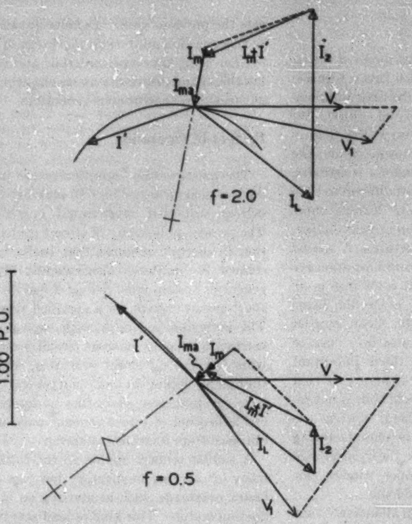


Fig. 4. Phasor diagrams, 80-per-cent power factor lagging load

load before the decay of rotor flux takes place, excitation is lost. As mentioned in the previous paragraph, only a short time is available to accomplish the transfer. In addition, the newly connected lagging load must be reduced quickly to a value within the system rating.

The wave form of the reactive current drawn by the main magnetic amplifier is poor, and the resultant nonsinusoidal machine current introduces wave form distortion into the output voltage. There is no easy way to eliminate this distortion.

Variable speed operation is possible. Operation at a regulated frequency is possible if the speed of the rotor can be adjusted to account for the slip. Parallel operation at a regulated frequency, with real and reactive load compensation, could be accomplished using the present hydraulic drive motors with but little change in the associated load-dividing circuits. Because of the lack of overload capacity already discussed, greater difficulty in "synchronizing" generators would be expected than is encountered with synchronous machines.

Some indications of the types of instability to be expected in a closed regulating system, and the way in which system design might be directed to minimize the corrections required in the regulating loop, are discussed in the section "Experimental Work."

Table I. System Data

Quantity	On System Base	On Machine Base
X_{C1}	0.56	
X_{C2}	1.60	
X_T	0.40	0.198
X_m	9.00	4.45
R_1	0.08	0.0396
R_2'	0.08	0.0396

All values are in per unit at base frequency

Possible Applications

The limitations already enumerated seem to prohibit the use of the induction generator for continuous duty in larger ratings, as now required in commercial or piloted military aircraft. However, the machine itself is adaptable to vaporization cooling and a high degree of miniaturization. Its manufacturing advantages, those of reduced cost and ready adaptability to mass production methods, indicate that it is better suited to quantity production than the synchronous generator. It appears that the induction generating system may be most useful in comparatively low-power, single unit, miniature ratings built in considerable quantity, where load circuit protection by limiters is not practiced.

Equivalent Circuit and System Design

The equivalent circuit of the system is shown in Fig. 2. The induction machine is represented by its simplified equivalent network having two parallel branches. The main magnetic amplifier is represented as a reactive element drawing a controllable current. Both shunt and

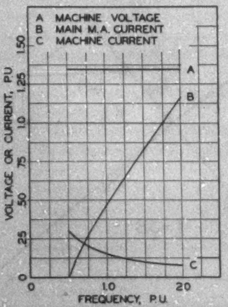


Fig. 5. Calculated performance, no load

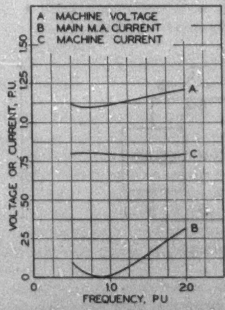


Fig. 6. Calculated performance, full load, 80-per-cent power factor lagging

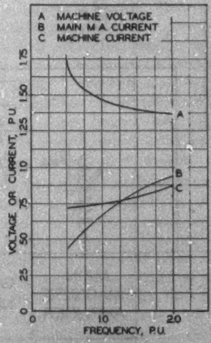


Fig. 7. Calculated performance, full load, unity power factor

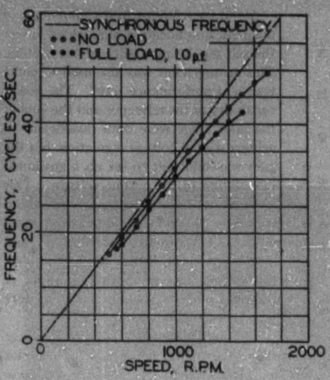


Fig. 8 (left). Frequency as a function of speed at various loads

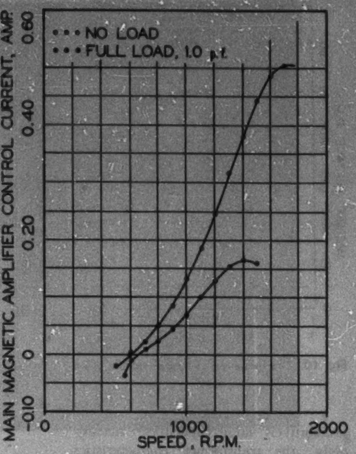


Fig. 9 (right). Main magnetic amplifier control current characteristics as a function of speed and load

series capacitors are included. All quantities are in per unit referred to the base of design center frequency, rated load voltage per phase star, and rated load current per terminal.

The circuit equations may be simplified to the form

$$V_1=V\left(1+\frac{X_{c1}}{X_{c2}}\right)-j\left(\frac{X_{c1}}{f}\right)I_L \qquad (1)$$

$$0=I_{ma}+I'+I_m+I_L+j\left(\frac{f}{X_{c2}}\right)V \qquad (2)$$

where all voltages and currents are phasors. One other equation is necessary for determination of V_1, I_{ma}, and s. This third equation states that the shaft input must be equal to the losses plus the electrical output. In other words, the total power delivered to the equivalent circuit from the point where V_1 is measured must be zero. The power equation is

$$I'^2\left(\frac{R_2'}{s}+R_1\right)+VI_L\cos\theta_L=0 \qquad (3)$$

An optimum choice for X_{c1} and X_{c2} minimizes the range of magnetic amplifier currents required for operation at full load over an extended frequency range. At low frequency, the presence of the series capacitor introduces a phase shift between the machine and the load such that the load current looks less reactive to the machine; the amount of reactive current needed from the shunt capacitor is considerably reduced. On the other hand, the generator should operate near its maximum power capability at high frequency, so that the extra reactive current provided by the shunt capacitor can be absorbed by the machine and need not be taken by the magnetic amplifier.

The condition that the induction generator operate at maximum power capability can be expressed analytically by the equation

$$VI_L\cos\theta_L=\frac{V_1^2}{2fX_T} \qquad (4)$$

where X_T is the sum of stator and rotor leakage reactances at base frequency referred to the stator. This equation is a statement that at rated load the output power component of machine current is just equal to the radius in the machine circle diagram. Substitution of this condition into equations 1 and 2 yields

$$\left(\frac{X_{c1}}{f}\right)^2-\left(2\frac{V}{I_L}\sin\theta_L\right)\left(1+\frac{X_{c1}}{X_{c2}}\right)\left(\frac{X_{c1}}{f}\right)+\left(\frac{V}{I_L}\right)^2\left(1+\frac{X_{c1}}{X_{c2}}\right)^2-2f\frac{V}{I_L}X_T\cos\theta_L=0 \qquad (5)$$

This equation gives a relationship between the ratio (X_{c1}/X_{c2}) and X_{c1} itself, with frequency and load as parameters. If it is assumed that $f=2.0$ per unit (twice design center frequency) and $X_T=0.4$ per unit at base frequency, it is found that X_{c1}/X_{c2} should be about 0.3 and X_{c1} should be about 0.6 per unit at base frequency. This combination is adequate to handle rated kva at any power factor between 0.8 lagging and unity.

Two sample phasor diagrams are shown in Fig. 3 for unity power factor, full load, and in Fig. 4 for 80 per cent power factor lagging, full load. Computations made in this way lead to performance data summarized in Figs. 5, 6, and 7. The system data assumed for these computations are given in Table I.

The highest machine voltage is 1.76 per unit, at 0.5 per-unit frequency and full load, unity power factor. The highest machine current is 0.87 per unit, at 2.0 per-unit frequency at full load, unity power factor. The per-unit base for machine impedances referred to the machine rating is 1.76/0.87 = 2.02 times the system per unit base. Machine impedances on their own base are shown in Table I. It is found that the machine assumed for sample calculations has lower than normal leakage reactance. The proper combination of series and shunt capacitance for other machine constants can be determined by trial calculations similar to those reported here.

Experimental Work

An induction generator system with series and shunt capacitors has been tested in the laboratory with results in general agreement with the calculated performance data. The machine tested was an industrial type of induction motor rated at 2.5 horsepower, 220 volts, 60 cycles, 6.4 amperes, 1,750 rpm. Rated output voltage was taken to be 30 volts, in order to maintain rated machine volts per cycle at an expected minimum frequency of 15 cycles with unity power factor, full load. Base frequency was taken to be 30 cycles.

The series capacitor reactance was 0.90-ohm at base frequency, while the shunt capacitor reactance was 1.70 ohms per phase star at base frequency. These low reactances were obtained by connecting capacitors at the secondary of step-up transformers in order to avoid the use of

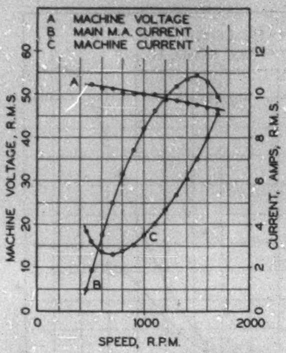


Fig. 10. Performance of model system at no load

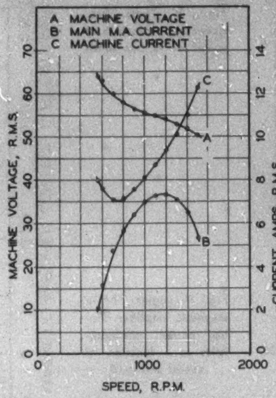


Fig. 11. Performance of model system at full load, unity power factor

prohibitively large capacitors. Accordingly, the transformer characteristics at the different frequencies acted to change the apparent capacitive reactance. A system designed for normal aircraft voltage and frequency would not have this difficulty since the capacitors could be connected directly.

The main magnetic amplifiers were controlled manually to obtain the steady-state performance data presented in Figs. 8 through 11. The lower limit of speed was set by the quiescent current of the main magnetic amplifiers, which increases rapidly at the lower frequencies.

The upper limit of speed was set by a stability consideration. It will be noted that in the region of higher speeds, an increase in speed requires a decrease in the main magnetic amplifier current to hold rated voltage, because of the transfer of reactive current to the machine. (This effect is most pronounced at full load; however, the high system losses in this laboratory model at no load cause the same behavior.) At some critical frequency the control current must be reduced as the speed (and frequency) increase further. Speed variations in this region set up an unstable oscillation because of the feedback through self-saturation of the main magnetic amplifier, and the result is loss of excitation. Use of somewhat less than 100 per cent self-saturation in the main magnetic amplifier could be expected to help control this type of instability should it appear in a closed loop regulating scheme.

A voltage reference and a pilot stage of magnetic amplification were added to test the self-regulating ability of the system. This arrangement was also used to make transient performance tests. The power capability of the reference and the gain of the pilot stage were both insufficient to achieve a good voltage regulation, but the gain was sufficient to set up a sustained modulation of the output voltage. This modulation was accentuated by the speed transients introduced into the prime mover. This type of modulation could be corrected in the usual way by the addition of feedback circuits to the input of the pilot stage, although this procedure was not carried out here

The modulation of output voltage shows the typical nonsinusoidal behavior associated with the variable response time of magnetic amplifiers However, this particular unstable behavior is not the same as the ferro-resonant oscillations which have been observed in series circuits containing capacitance and saturable reactors. In fact, the induction generator system with shunt and series capacitors can be thrown into ferro-resonance by the sudden application of an inductive load of low impedance. The subsequent transient dies out very slowly (if at all) and is characterized by very severe subharmonic oscillations. This type of ferro-resonant oscillation cannot be corrected by the usual feedback techniques, since an entirely different mode of operation exists in the magnetic amplifiers. Ferro-resonant oscillations have not been observed when the excitation is taken from shunt capacitors alone, but a wide frequency range is not practical unless series capacitors are used in combination with shunt

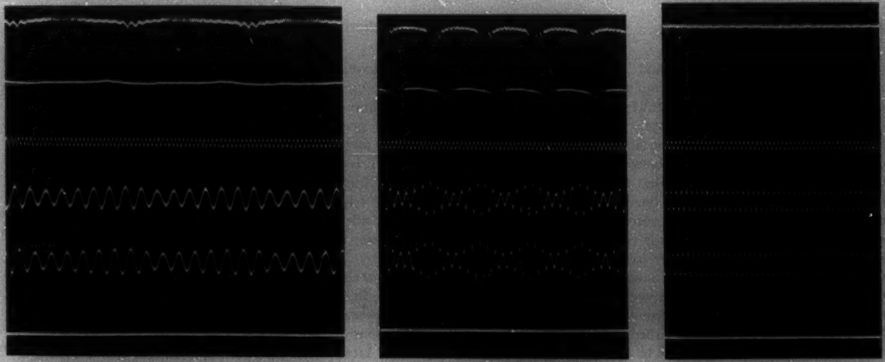
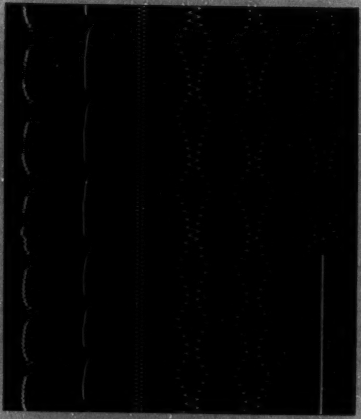
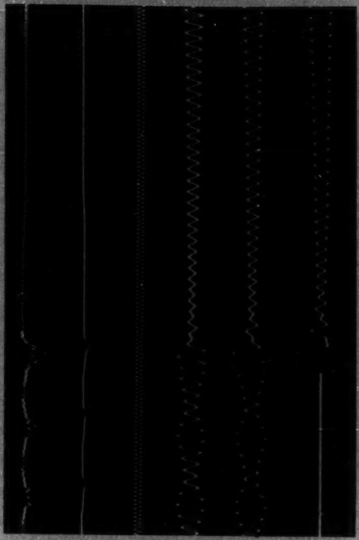
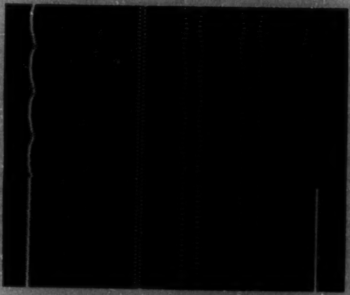

Serial 7 **Serial 8** **Serial 9**

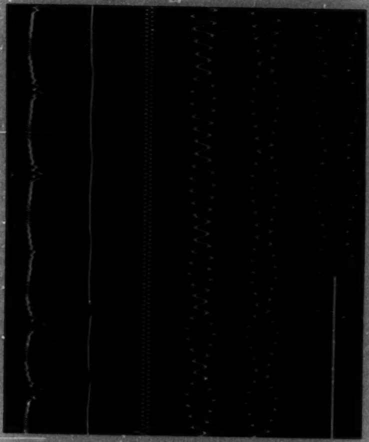
Fig. 12. Transient performance of model system. Oscillograms identified in Table II (Fig. 12 continued on page 293)

Fig. 12 cont.


Serial 10


Serial 11


Serial 12


Serial 13

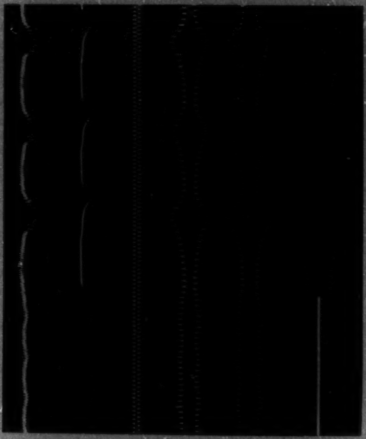

Serial 14

Table II. Identification of Oscillograms in Fig. 12

	Description
Serial	
7	no load, 500 rpm
8	no load, 1,000 rpm
9	no load, 1,800 rpm
10	balanced 0 power factor inductive load applied, 1,000 rpm
11	balanced 0 power factor inductive load applied, 700 rpm
12	balanced 0 power factor inductive load applied, 1,600 rpm
13	balanced resistive load applied, 700 rpm
14	balanced resistive load applied, 1,600 rpm
Channel	
1 (top)	pilot magnetic amplifier control current
2	main magnetic amplifier control current
3	60-cycle timing
4	line voltage V_{ab}
5	line voltage V_{ca}
6 (bottom)	Δ phase current I_{bc}

capacitors. Considerable further work is necessary to determine the frequency range over which the system can operate with minimum danger of ferro-resonant instability. The problem is complicated by the fact that each frequency range has its own optimum proportion between series and shunt capacitors to minimize the main magnetic amplifier weight.

The performance of the self-regulated system is illustrated by the oscillograms of Fig. 12. Identification of the various oscillograms is given in Table II.

Conclusions

This study of the self-excited induction generator with magnetic amplifier control, although by no means complete, has demonstrated that the generating system is workable. The main problems requiring further extensive investigation are the over-all weight in comparison with the synchronous machine, the optimum choice of main magnetic amplifier and capacitor combination to minimize system weight and at the same time maintain inherent stability of control over a wide speed range, and the prevention of ferroresonant behavior during load transients. The study has indicated that all of these problems have a practical solution. A reasonable field of application for the system has been indicated after a discussion of the design advantages and operational limitations.

The experimental work undertaken so far has been a kind of demonstration. The constants assumed for analysis were representative only; they were not those associated with any particular machine. Also no attempt was made in the analysis to account for the effects of transformer losses and exciting currents, which were peculiar to the laboratory model. Thus the agreement between analysis and experiment is of a general nature only, and offers encouragement but not numerical coincidence. The oscillating behavior noted in some of the oscillograms is peculiar to the model chosen for demonstration and does not represent a universal defect.

The limitations mentioned in the body of the paper, together with those difficulties described in the section on experimental work, indicate that probably about 2 or 3 kva is an upper limit of rating for this induction generator system. Of course this figure is speculative and must await a comprehensive weight analysis.

◆

Discussion

Elvin D. Lytle, Jack B. Caldwell, and **Robert C. Byloff** (The Garrett Corporation, Los Angeles, Calif.): The statement, "The most serious limitation of the induction generator system is its inability to deliver sustained short-circuit current" requires additional discussion. As a matter of fact, the system proposed in the paper will deliver sustained short-circuit current. It can be seen in Fig. 2 that if a short circuit is put across the load terminals, the series capacitor is connected across the terminals of the generator to furnish excitation. The excitation current thus produced will flow through the applied short circuit. In the system discussed in the paper, assuming the per-unit values given are equivalent line-to-neutral values, the series capacitor alone has three times the capacity of the shunt capacitor and four times the effective capacity of the series and shunt capacitors together. Since the series capacitor will always be three or more times as large as the shunt capacitors in any practical system, any system using shunt and series excitation capacitors can be made to supply sufficient sustained short-circuit current for fault-clearing purposes. This general analysis has been confirmed by extensive laboratory tests performed by the writers.

Hence, contrary to the statement in the paper, the induction generator system proposed actually acts more like a compound wound d-c generator than it does a shunt generator. Performance similar to a d-c shunt generator would be obtained if only shunt excitation capacitors were used between the terminals of an induction generator. But even in this case, the fault-clearing ability of the induction generator should be given further consideration.

Although it is true that a simple shunt capacitor-excited induction generator cannot deliver sustained short-circuit current, it should be remembered that the requirement of the generator user is not for short-circuit current in itself, but for fault-clearing ability. Actually, contrary to the statement of the author that ". . . the induction generator time constant is so short that no appreciable heating of a fuse element may be obtained from its transient current," it has been the experience of the writers that the transient decay currents possess considerable fault-clearing ability.

One of the inherent features of most induction machines is its low internal impedance. As a result, experience shows that the amplitude of the current, upon application of a short circuit, reaches values of as high as 25 times the rated load current of the machine. For example, one 15-kva, 4-pole, 400-cycle machine (rated current of 43.4 amperes at 200 volts line to line) frequently produces transient currents of 1,100 amperes (rms) upon application of a short circuit. It has been found that these high values of transient current will open thermal-type circuit breakers with current ratings as high as one-third of the full load current of the generator. Similar results are obtained with fuses. Fault-clearing times with magnetic circuit breakers of current rating equal to 100 per cent of machine-rated current have been measured at less than 0.01 second.

The author also voiced apprehension concerning the ability of an induction generator to build up voltage after an inactive period of several days. Out of several hundred induction generators of five different models manufactured at the writers' plant, none has ever failed to build up voltage. In fact, newly assembled machines will build up voltage the first time they are operated without benefit of any magnetizing operation. Machines which have been idle for periods of several years build up voltage without fail when put into service again. It has been found, however, that there is a critical value of capacity and speed below which an induction generator will not build up. This critical value has been found to be affected slightly by disassembly of the machine, by the load connected to the machine when stopped, and by the applied load when starting. However, a reasonable value of capacity will always permit voltage build-up. The use of insufficient excitation capacity is the most probable explanation of the author's experience in finding that some generators with low residual magnetism would not build up.

It would be of extreme interest and value to see what the practical and competitive ratings of regulated induction generator systems might be with speed ranges less than four to one. In view of the very marked trend of airframe manufacturers to specify electrical systems with ever-narrowing frequency ranges, and in view of the constant improvements in the field of constant speed drives, there would be considerable value in knowing the limits of practicability in systems with frequency ranges of ±20 and ±5 per cent. Such an analysis would undoubtedly show a much wider range of usefulness for the induction generator.

R. A. Larson (Westinghouse Electric Corporation, Lima, Ohio): A possible control system for an induction generator other than the one outlined in this paper could consist of an inductor-type synchronous condenser connected in parallel with the induction generator. The synchronous condenser would need sufficient capacity to compensate for the most severe inductive loading within the operating load and speed range plus the required induction generator excitation. Shunt capacitors would be required; however, their capacity need be only sufficient to build up the system, i.e., the induction generator and the synchronous condenser which would necessarily have to start as an induction motor.

Voltage control would be accomplished by varying the field excitation of the synchronous condenser by means of either a carbon-pile or Magamp regulator using the rectified output of the induction generator as the source of field supply. An induction generator, thus self-excited and controlled, would have essentially the same inherent advantages and limitations as outlined in the paper, except for its outstanding advantage: the wave form of the output voltage would be improved.

The inductor-type synchronous condenser would be as rugged as an induction generator, since no brushes are required and no rotor windings exist. The output per-unit weight of the inductor-type synchronous machine would be substantially less than that of the induction generator. The resulting net output per-unit weight of an induction generator, synchronous condenser, and regulator would be low in comparison with the more common synchronous-generator systems.

S. A. Monfort (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The author has presented an interesting application of magnetic amplifiers in the control of voltage on induction generators. We must agree however, that the resulting output voltage wave form is rather poor, and it is not expected that such distortion as indicated on the test oscillograms would be tolerated on an electric utility system. It is perhaps possible, with filter circuits, to improve this condition. However, the possibility that an effective regulator employing magnetic amplifiers and necessary modifications for improvement of wave form could be produced at a cost to make it economically justifiable for use on central station size induction generators is somewhat questionable.

As is pointed out, there are certain design features of the induction generator which would seem to make this unit more desirable than synchronous machines in systems of rather small capacity, such as aircraft service. The use of induction generators has also been economically justified in a few central Station installations. There is no doubt that there are certain economical advantages of the induction machine in the small capacities as described by the author. However, as the rating of the induction generator increases, mechanical and electrical requirements dictate that the machine be physically oversize, as compared with synchronous machine, which in turn leads to lower power factor and efficiency, as well as higher cost. In practically all cases this additional cost has been found to offset the saving by elimination of an exciter.

When induction generators receive their excitation from shunt capacitors, there is always the possibility, when the generator and capacitors as a unit are disconnected from the system, of self-excitation taking place and causing a build-up of generator voltage of dangerous proportions, especially in the case of water wheel generators because of the high overspeeds involved.

H. M. McConnell: The author is grateful to Mr. Lytle, Mr. Caldwell, and Mr. Byloff for the information they have added in regard to the fault clearing capability of the induction generator, and for their careful attention to the inconsistency in the paper on that subject. It is agreed that the system with series and shunt capacitors can be made self-excited under short-circuit conditions with behavior similar to the compound d-c generator.

This discussion also raises the question of optimum design for a constant-frequency system. Shunt capacitor excitation is indicated for this condition, and the author's work began with the magnetic-amplifier control of such a system in both constant and limited variable-frequency service. However, the advantages of compound-capacitor excitation in fault-clearing capability indicate that the compound excitation should be investigated for constant-frequency applications.

With regard to the build-up problem, it should be mentioned that the discussers undoubtedly have reference to their experience with generators having voltage control by inherent saturation. In such a system it is extremely unlikely that zero power factor reactive loads are connected to the machine while starting. However the magnetic-amplifier voltage control proposed in the paper constitutes such a load, and the build-up problem is more severe if the quiescent current of the main magnetic amplifier is excessive.

Mr. Larson's proposal to provide excitation by means of a synchronous condenser is an extension of the central station practice of providing excitation for induction generators as part of the reactive load on the synchronous generators located elsewhere. It is true that large blocks of controlled capacitive kva are more economically provided by synchronous machines than by static capacitors. However, the author sees some difficulty in constructing an inductor-type polyphase synchronous machine. The only possibility seems to be three separate stator structures arranged axially, with mechanical phase displacement. Also there is the problem of maintaining synchronism with the induction generator frequency, which will vary suddenly according to the kilowatt demand, even though the shaft speed of the induction generator is constant. Resulting power-angle swings of the synchronous machine might well cause system instability.

The author agrees with Mr. Monfort that the system proposed in the paper will find very little application to central station service. The induction generator is itself quite practical in ratings of, say, a few hundred kva at an unattended station operated in parallel with synchronous generation, where the voltage is fixed by the system. However, the voltage control and excitation proposed in the paper would be quite absurd in such ratings.

Application of Large A-C Motors and Controls in Refineries and Pipe Lines

G. L. OSCARSON
MEMBER AIEE

MOTORS of 250 horsepower (hp) and larger, as used in oil refineries, may drive reciprocating compressors, centrifugal compressors, or centrifugal pumps. They may be located in areas where explosive hazards exist in varying degree, where corrosive or abrasive conditions prevail, or they may operate in normal ambient conditions.

Large motors, operating on trunk pipe lines, perform in moderately hazardous conditions, whether in crude-oil or products liquid lines or in natural gas lines. They are frequently located in areas where abrasive dusts are present in the air.

Consideration of these factors has led to the development of various degrees of protection, with respect to climatic, corrosive, or abrasive conditions and to explosive or electrical hazards. The purpose of this paper is to evaluate the influence of these factors, as well as of general load characteristics, in so far as they affect the application of motors and controls rated 250 hp or larger.

Refinery Applications

LOAD CHARACTERISTICS

Reciprocating compressors may handle air, refrigerants, or products gases. They are unloaded when starting, requiring not over 40-per-cent starting torque, and having full-speed no-load torque requirement of not more than 30 per cent of full-load torque. Speeds range from 200 to 600 rpm with unit ratings as large as 2,000 hp.

Centrifugal compressors handle air or products gases and operate at speeds of 3,600 to 10,000 rpm with requirements to 5,000 hp. These units may readily be unloaded for starting by closing inlet and discharge dampers. The starting torque required is only enough to overcome static friction. The full-speed no-load torque may range from 35 to 60 per cent of full-load torque. By venting the compressor casing to atmosphere the full speed, no-load torque may frequently be lowered to approximately 10 per cent. At 3,600 rpm these compressors are direct-coupled to the driving motors; at higher speeds they are driven through gearbox speed increasers. Fig. 1 illustrates a 3,000-hp, 1,780-rpm forced-ventilated motor, totally enclosed, driving a centrifugal blower through a speed increaser.

Centrifugal pumps for refinery service may operate at speeds ranging from 720 to 3,600 rpm depending on the hydraulic conditions. They are direct-coupled to the driving motors. Starting torque required is only that necessary to overcome static friction; full-speed no-load torque requirement with closed gate (closed discharge valve), is usually from 50 to 60 per cent of full-load torque.

MOTOR CHARACTERISTICS

In some refineries it is the practice to consider the separate divisions as process units with load requirements up to 5,000 kva each, with high-voltage distribution at 13,200 volts and with individual transformer banks for each process unit. In such cases, motors of 250 hp and higher should be 2,300 or 4,000 volts from the standpoint of overall first cost, operating characteristics, and trouble-free operation. A possible exception would be the use of 13,200 volts for 3,600-rpm synchronous motors of 5,000 hp or larger, which are essentially of turbine-generator construction.

We may assume that the drives just discussed are all fundamentally constant-speed applications and may be driven either by squirrel-cage motors or synchronous motors; the selection being made on such factors as simplicity, first cost, efficiency, power factor, and mechanical ruggedness.

Reciprocating compressors, requiring motors of 250 hp or larger, are usually driven by engine-type motors; the motor rotor being mounted on the compressor shaft and the stator supported either on the compressor frame or on soleplates set in the foundation. Synchronous motors are generally used because of their superior performance at the relatively low speeds involved.

Squirrel-cage 3,600-rpm motors, through 5,000 hp, have acceptable characteristics for centrifugal compressor drive and are usually preferred. Above 5,000 hp the forged-rotor turbo type of construction of synchronous motors presents some definite advantages. However pull-in torque values over 15 per cent can be obtained only at the expense of high current values.

At 1,800 rpm, squirrel-cage motors have been applied through 3,500 hp, having quite acceptable efficiencies and power factors at that speed. At motor speeds of 1,200 rpm or less, synchronous motors have significant advantages of efficiency and power factor and should be given first consideration. They have been applied to centrifugal compressors in ratings through 4,500 hp at 1,200 rpm and 6,000 hp at 900 rpm. Centrifugal pumps at 1,800 or 3,600 rpm will normally be driven by squirrel-cage motors. In lower speeds the efficiency and power factor advantages of synchronous motors should be given due consideration. The foregoing information is shown in Table I.

PROTECTIVE ENCLOSURES—INDOOR INSTALLATIONS

For indoor installation in nonexplosive and noncorrosive areas, standard open or drip-proof motors are used. When concentration of an explosive mixture adjacent to the motor is possible, it is necessary to guard against a possible igniting condition. This can be done by using explosion-proof motors, which are not airtight but which will withstand an internal explosion without propagation of flame outside the motor casing. These motors are impracticable to build in the larger sizes and they always have a high first cost. Maintenance is rather difficult because of the nature of the enclosure.

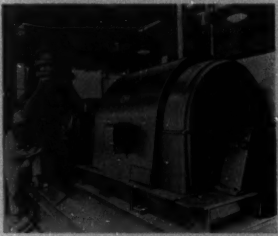


Fig. 1. 3,000-hp 1,770-rpm squirrel-cage motor driving high-speed blower through speed increaser

---

Paper 54-526, recommended by the AIEE Petroleum Industry Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Petroleum Industry Conference, Tulsa, Okla., September 27–29, 1954. Manuscript submitted June 16, 1954; made available for printing September 15, 1954.

G. L. OSCARSON is with the Electric Machinery Manufacturing Company, Minneapolis, Minn.

Table I. Recommended Motor Type for Various Applications

Application	Horsepower	Motor Rpm	Motor Type
Reciprocating compressor	250– 3,000	600– 300	synchronous
Centrifugal compressor	250– 5,000	3,600	squirrel-cage
	5,500–12,500	3,600	synchronous
	250– 3,500	1,800	squirrel-cage
	250– 6,000	1,200– 900	synchronous
Centrifugal pump	250– 3,000	1,800–3,600	squirrel-cage
	250– 3,000	1,200– 720	synchronous

One solution to the problem of keeping corrosive or explosives gases out of the motor is to use inert gas or instrument air, under slight pressure, as the cooling medium. A water-cooled heat exchanger is used to carry away the heat. Internal pressure is maintained at about 2 inches of water. If the pressure drops to 1 inch an alarm is sounded, and at 1/2 inch the motor is shut down. Oil shaft seals may be used, on either high- or low-speed units, to limit the leakage per day per seal to an insignificant amount.

Totally enclosed fan-cooled motors are fundamentally similar to explosion-proof motors except that they may not be used in hazardous explosive areas.

In some cases satisfactory cooling air is available within a reasonable distance of the motor location. The motors may then be built with totally enclosed forced-ventilated housings, with cooling air supplied by a motor-driven blower. Interlocking is provided so the blower motor must be started first and the system purged before the main motor can be started; also the main motor is shut down immediately on failure of cooling air supply. It is important that the entire intake duct be kept under positive pressure to prevent encroachment of explosive or corrosive gases in case of a leak. In some areas intake filters must be provided to exclude abrasive dust particles.

Protective Enclosures—Outdoor Installations

The recent trend is toward outdoor installation whenever possible of centrifugal compressors and pumps, their associated motors, and in some cases the motor controls. The totally enclosed fan-cooled motor has long been used for outdoor installation, except in areas where explosive gases may prevail, in which case the explosion-proof construction is required. One disadvantage to such motors exposed to direct sunshine is that a sudden rain shower will result in a sharp drop in temperature and the breathing in of moisture-laden air. Drains may be impossible to use or may be ineffective, in which case several inches of water may collect in the motor housing.

A recent development is the use of the weather-protected motor, which is an open-type motor with the ventilating passages designed to minimize entrance of air-borne rain, snow, or abrasive particles. Where exposure to air-borne abrasive particles is especially severe, intake filters should be provided.

All outdoor motors should be provided with space heaters which are energized automatically when the motor is shut down. This will serve to minimize condensation of moisture. Where explosive vapors may be present the surface temperature of space heaters should be limited to 200 degrees centigrade.

Liquid Pipe-Line Applications

Load Characteristics

All major electrified crude-oil or products lines built in the United States within the past 15 years have used centrifugal pumps. Speeds will usually be 3,600 rpm, with 1,800 rpm being used on some large ratings. The general characteristics will be similar to pump applications in refineries except that the motors will more often be controlled automatically or started by remote control, will be subjected to more severe voltage dips during starting, and may be subjected to more extreme climatic conditions. Any conditions of load or operation differing from those previously discussed will be covered in detail in this section.

Motor ratings in large pipe lines, will range through 3,000 hp at 1,800 rpm and 1,500 hp at 3,600 rpm. In attended stations the pumps will be started closed-gate and the torque required at full speed will be about 50 per cent. In automatic or remote-controlled stations it may be desirable to initiate opening of the discharge valve before the pump is started, to prevent any wedging action due to

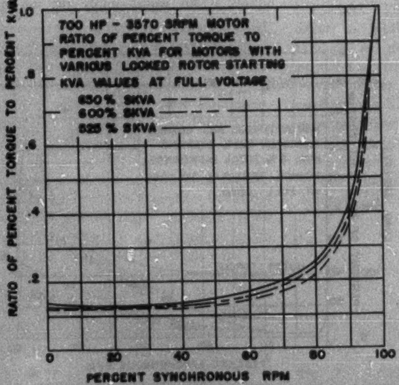


Fig. 2. Curve showing relationship of torque to kva for various locked-rotor kva input designs

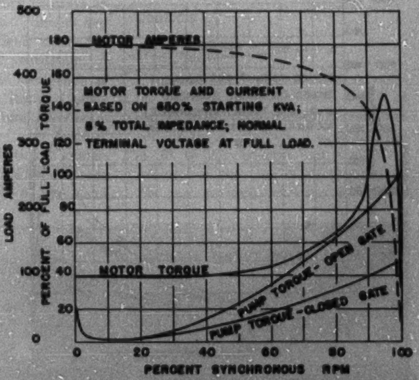


Fig. 3. Motor torque and amperes versus speed based on 8-per-cent total-system impedance, 650-per-cent locked-rotor kva

pressure build-up. Owing to the rather slow action of the valve-opening mechanism, the motor will normally be up to speed before the valve is wide open. However, if the pump is discharging into an open line its characteristic is such that open-gate torques can be approached even with the discharge valve still partially closed. It may be assumed that, during acceleration, the pump torque will vary as the square of the speed.

MOTOR CHARACTERISTICS

Pipe-line motors are usually supplied with energy from individual substations. Transmission voltages range from 11,000 to 110,000 volts. The utilization voltage is usually either 2,300 or 4,000 volts; transformer capacity of approximately 1 kva per hp is commonly provided. The motors are started across the line (full voltage) whenever possible, so the step-down transformers usually have at least 6-per-cent impedance to reduce current inrush and to minimize disturbance to other power users. Total impedance, to a constant voltage source, often approaches 10 per cent with transformer rating as a base.

These motors are virtually always of the squirrel-cage type and the torques developed will vary as the square of the applied voltage. The starting kva, on full voltage, will normally be about 650 per cent of full-load kva. This will result in an appreciable voltage dip on a high-impedance supply system and the torques will be reduced accordingly.

In some cases, motors having less than 650 per cent of full-load starting kva have been applied to pipe-line pumps. Assuming that full-load slip, full-load efficiency, and full-load power factor values are fixed, the starting kva and accelerating torques are bound together by the leakage reactance. Fig. 2 illustrates the close relationship between torque and kva, with starting kva values in the range of 525 per cent to 650 per cent of full-load kva, of typical pipe-line motors.

As mentioned previously, a substantial voltage dip can be expected during starting most pipe-line motors. It would appear that a low-starting kva motor would be preferable inasmuch as it would entail the least impedance drop. However, that is largely offset by the lower accelerating torque developed.

Figs. 3, 4, and 5 illustrate the accelerating torques developed by 700-hp, 3,600-rpm motors having starting kva values of 650 per cent, 600 per cent and 525 per cent respectively of full-load kva. They were assumed connected to a system having a total impedance, to a regulated voltage point, of 8 per cent with motor full-load kva as a base, the transformer taps arranged to provide rated motor terminal voltage at full-load current. It will be noted that if open-gate torque conditions are assumed, the motors involved would accelerate to full speed. However, this study points up the fact that 10-per-cent impedance would have resulted in sufficient decrease in torque to prevent the motor from accelerating to full speed under open-gate conditions.

A reasonable requirement would be that the terminal voltage should recover to 80 per cent of normal at a current value corresponding to motor input at 70-per-cent speed, and that the motor should develop at least 55-per-cent torque at 70-per-cent speed and 65-per-cent torque at 80-per-cent speed with 80-per-cent applied voltage. If such conditions are realized there should be no acceleration difficulty experienced in automatic or remotely controlled stations.

Some discussion as to accelerating time may be in order. Referring to Fig. 5, assume that the pump will be started closed-gate. The net accelerating torque at any one speed will be the pump torque at that speed subtracted from the motor torque. Obtain these values at 10-per-cent speed intervals, from zero to full speed, and average them; it is found that the average net accelerating torque is 35 per cent, then motor full-load torque = hp × 5,250/full-load rpm = 700 × 5,250/3,565 = 1,390 pound-feet. Average net accelerating torque is then 1,390 × 0.35 = 487 pound-feet.

$$t=\frac{WK^2\times \text{rpm}}{308\times Tn}$$

where

t = time in seconds
WK^2 = total WK^2 of motor and pump
Tn = average net accelerating torque

The WK^2 value must be obtained from motor and pump manufacturers. For 700-hp 3,600-rpm units this value will be approximately 400 pound-feet.[2] Then accelerating time is

$$t=\frac{400\times 3{,}565}{308\times 487}=9.5 \text{ seconds}$$

According to the National Electric Code,[1] Class 1, Division 1 locations would usually include such locations as pump rooms for volatile flammable liquids. This would imply the use of

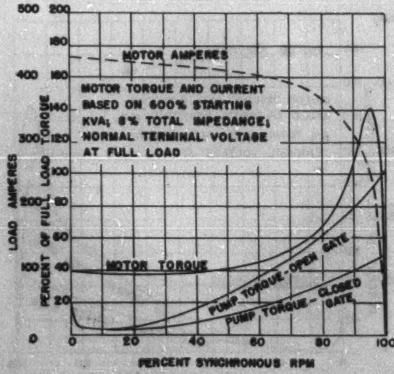


Fig. 4. Motor torque and amperes versus speed based on 8-per-cent total-system impedance, 600-per-cent locked-rotor kva

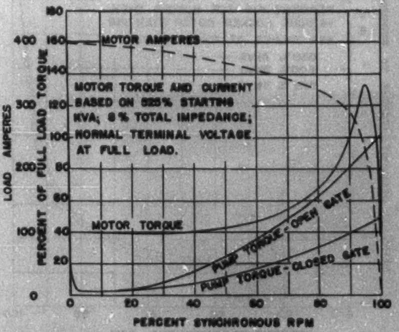


Fig. 5. Motor torques and amperes versus speed based on 8-per-cent total-system impedance, 525-per-cent locked-rotor kva

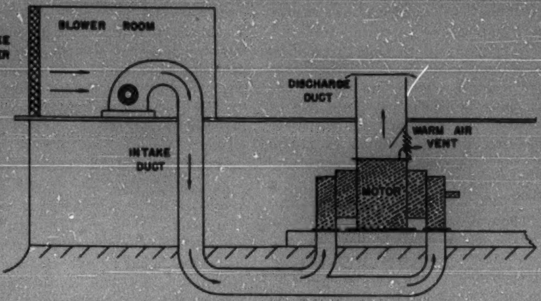


Fig. 6. Diagram illustrating installation of 3,000-hp 1,780-rpm totally enclosed forced ventilated squirrel-cage motor in pipe-line station

explosion-proof motors. However the petroleum industry has widely accepted the use of totally enclosed forced-ventilated motors for this application. In the early days of pipe lining, many open-type motors were installed in pump rooms. There are very few reported cases of trouble resulting from the hazards involved. Many installations have been made using open-type motors with fire walls between the motors and the pumps. However, this construction presents some operating and maintenance problems.

Present practice in indoor installations is to install the motors in the pump rooms. Relatively small installations frequently use explosion-proof motors. However, these present maintenance problems and first cost are high. As a result, particularly in the larger sizes, many totally enclosed forced-ventilated motors are being used. The cooling air is usually filtered and the intake duct kept under pressure by means of a blower. Fig. 6 shows, in cross section, a typical totally enclosed, forced-ventilated motor installed in a pipe-line station. Motor-discharge air may be used for room heating.

In pipe lines, as in other equipment, there is a growing tendency toward outdoor installation. Explosion-proof motors are frequently used in the smaller ratings, to protect against both the elements and explosive hazards. In ratings of 1,000 hp or larger, explosion proof motors are rarely used because of manufacturing difficulties and high cost.

Outdoor weather-protected motors are becoming increasingly popular. Some authorities feel that the possibility of explosive concentration in outdoor installations involving crude or products lines is extremely remote; and further that the possibility of such a condition simultaneous with an igniting condition in the motor is too remote to constitute an appreciable hazard. In dusty areas weather-protected motors should be provided with intake filters. Fig. 7 illustrates two 1,000-hp 3,550-rpm weather-protected motors, with intake filters, installed in a pipe-line station.

In some locations there is a strong possibility of sufficient snowfall to blanket a motor completely if it is not running. As a result, the intake and discharge ducts may become blocked, resulting in virtually no air circulation. If such a motor is started automatically, or by remote control, there may be no knowledge of the necessity of unblocking the openings. For such locations a desirable feature is a by-pass opening which will permit discharge air, if the discharge opening is blocked, to flow into the intake duct. Such a by-pass would be opened with the advent of cool weather and the resultant recirculation would be insignificant unless the discharge opening is blocked. The motor then becomes, in effect, a totally enclosed motor and remains so until sufficient snow has melted to permit normal ventilation.

The matter of proper lubrication of outdoor pipe-line motors, installed where extremely low temperatures are occasionally encountered, is of some concern. One solution is the installation of heaters in the bearings permitting, winter and summer, the use of oil of the same viscosity. However, a power outage, permitting the oil to stiffen, could create a real hazard. Any attempt to start the motor by remote control, soon after power restoration, could result in almost immediate bearing failure; bearing temperature relays probably being ineffective under those circumstances. A preferable procedure would be to use a heavy oil in the summer and to use a winter oil which has acceptable viscosity even at the lowest probable temperature. Oil change at the corresponding 6-month intervals would prevent the possibility of the development of excessive acidic conditions. Large high-speed pipe-line motors should have unidirectional blowers for quiet operation and efficiency. One method of obtaining whichever direction of rotation may be required in changing locations is to build these motors with double-shaft extensions. The unused extension is then capped to prevent its being a hazard to personnel.

Gas Pipe Lines

Large motors for gas pipe lines would be limited to centrifugal compressor drive. Motor ratings in excess of 10,000 hp have been considered for this duty. Ratings through 5,000 hp at 3,600 rpm or 3,500 hp at 1,800 rpm will usually be squirrel-cage. Motors at speeds of 720, 900, or 1,200 rpm, or larger than 5,000 hp at 3,600 rpm, will normally be of the synchronous type.

The same general considerations apply in the use of motors and their associated controls for gas pipe-line installations as have been discussed previously for refinery service.

Controls

STARTING VOLTAGE

In general, any of these motors, except the 2-pole synchronous ones, should be started on full voltage unless the resultant voltage dip is excessive. Occasional dips of 20 to 25 per cent can usually be tolerated if they are of short duration. The torques then developed will suffice

Fig. 7. 1,000-hp 3,550-rpm weather-protected motors in pipe-line station

for normal starting conditions. A 2-pole synchronous motor has a locked-rotor input of approximately 1,000 per cent. For this reason, and because of their size, these motors are always started on reduced voltage. Even at 96-per-cent speed the current input is so high that the motor should also be synchronized on reduced voltage if possible. A neutral reactor, short-circuited for full-voltage condition, often works out advantageously for this application.

Fault Current Protection

The matter of protection in case of line faults is an important one. In liquid pipe-line stations, and in refineries with segregated process units, the transformer impedance will usually limit the fault kva, on motor voltages of 2,300 or 4,000, to less than 150,000. Such faults are readily controlled by current-limiting high-interrupting capacity fuses. For normal-starting kva motors these fuses are adequate for ratings through 1,250 hp at 2,300 volts or 2,250 hp at 4,000 volts.

If fuses are not applicable, a high-interrupting-capacity circuit breaker is used, serving both as a line switch and a fault-current interruptor. Even when not required for fault interruption, and in spite of higher first cost, drawout breakers are sometimes preferred where duplicate units are involved, because of ease of maintenance.

Overcurrent Protection

Thermal overcurrent relays, of bimetallic-metal or fusible-alloy type, provide basic overcurrent protection for these motors. Induction overcurrent relays may be used to provide more rapid tripping in the region of locked-rotor currents. Fig. 8 illustrates the coordination of a current-limiting fuse, induction overcurrent relays, and thermal overcurrent relays in protecting a large motor. Any region to the right of the dotted line will be protected by the current-limiting high-interrupting-capacity fuse. This assures that any current in excess of 1,000 per cent of rated motor current will be interrupted by the fuses. The induction relay will provide quick circuit opening in case of stalled rotor conditions, thus protecting the squirrel-cage winding, but will not operate on anything less than 150 per cent of full-load current. The thermal relay will provide protection on any continuous or repeated currents in excess of rated value.

Undervoltage Protection

In general, time delay undervoltage protection should be provided for induction motors. However, any voltage dip will result in a drop in motor speed and an increase in line current. If the motor is large enough so that any such current increase will be a burden on a system that may already be in distress, the voltage protection should be instantaneous. Synchronous motors should be instantaneously removed from the system if they drop out of step as the result of a voltage dip.

Additional Protection

Motors of 1,500 hp and larger should be provided with stator temperature detectors and differential protection. Bearing temperature relays are desirable on all motors driving pumps which handle flammable liquids. Phase failure and phase reversal protection is also frequently used on large motors. In addition, synchronous motors should be provided with excitation failure relays and out-of-step relays.

Relay Function

Thermal overcurrent relays stator temperature detectors and bearing temperature relays normally indicate gradually developing conditions. They may be used either to initiate warning signals for the operator or to shut the equipment down. All other relaying would indicate fault conditions calling for immediate shutdown.

Indoor Controls

Explosion-proof controls are not commonly available in the voltages and hp

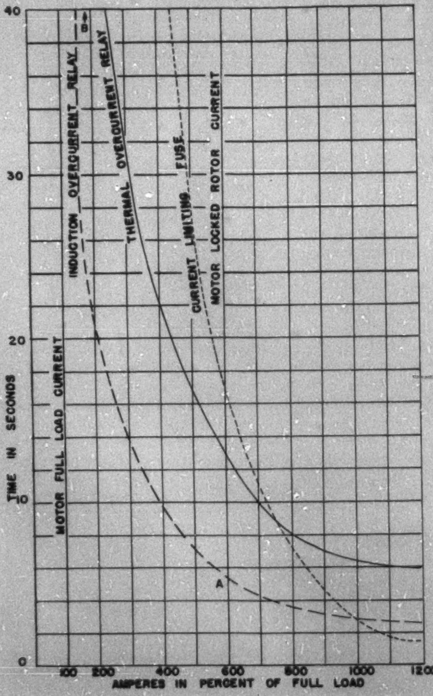


Fig. 8 (left). Curves showing co-ordination of high-interrupting capacity fuses, induction overcurrent relays, and thermal overcurrent relays for squirrel-cage motor protection

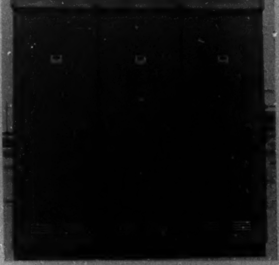


Fig. 9 (right). Outdoor installation of three 4,160-volt high - interrupting capacity fuse-type squirrel-cage motor controls ratings discussed here. Indoor controls are therefore of conventional construction, preferably located in areas where accumulation o' explosive gases is not possible. In some cases this cannot be done and the controls are then located in rooms which are pressurized by means of blowers located where uncontaminated air is available.

OUTDOOR CONTROLS

Outdoor types of controls are available for most ratings and are frequently used with outdoor-type motors. Fig. 9 shows high-interrupting-capacity fuse-type full-voltage controls as used with outdoor motors in a pipe-line station. These controls have screened openings to prevent entrance of small birds and rodents.

Occasionally some unusual requirements apply to outdoor controls. For instance, the use of spring-loaded interlocks has proved unsatisfactory owing to corrosion of sliding parts and failure of the spring-operated mechanisms. Positive operating means are therefore required. All control circuits, push-button stations, relays, or similar devices which are located in hazardous areas must be approved for the specific duty involved.

Exciters

Exciters for 3,600-rpm motors will be of the motor-generator type. If indoor control is used, the exciter sets may be put in the same location. With outdoor motors and controls the motor and generator may be of explosion-proof construction and not limited as to location.

For synchronous motors in the range of 600 to 1,200 rpm, located indoors in nonhazardous areas, the exciters may be standard direct-connected open-type units. In lower speeds, and for virtually all reciprocating compressor applications, direct-connected exciters are not practicable, and motor-generator sets will be used of either an explosion-proof or open type as conditions may dictate.

References

1. NATIONAL ELECTRICAL SAFETY CODE. National Bureau of Standards, Washington, D. C., section 5004a, 1951.

2. RELIABLE ELECTRIC POWER ESSENTIAL TO PROCESSING, T. L. Cappel, Jr. *The Petroleum Engineer*, Dallas, Tex., vol. 25, Nov. 1953, pp. 762–64.

3. A METHOD OF DETERMINING INDUCTON MOTOR SPEED-TORQUE-CURRENT CURVES FROM REDUCED VOLTAGE TESTS, R. F. Horrell, W. E. Wood. *AIEE Transactions*, vol. 73, pt. III, June 1954, pp. 670–75.

4. ELECTRICITY'S CONTRIBUTION TO PIPELINING, Merritt Hyde. *The Petroleum Engineer*, Dallas, Tex., vol. 26, Jan. 1954, pp. 605–22.

5. PIPE LINE APPLICATIONS OF ELECTRIC MOTORS, E. F. Greiwe. *The Petroleum Engineer*, Dallas, Tex., vol. 25, Nov. 1953, pp. 615–25.

No Discussion

Comments on Aircraft Switch Testing

T. R. STUELPNAGEL
ASSOCIATE MEMBER AIEE

J. P. DALLAS
MEMBER AIEE

Synopsis: It has been stated that the only real obstacle to push-button warfare is the design of an adequate switch. Unfortunately, aircraft switch design and application have lagged behind the need of the industry. One of the most effective ways to shorten this lag is to learn more about switches and switching problems. In this regard a scientific switch-testing procedure is proposed in this article. Consideration is given to factors that determine switch merit. These factors include insulation, contact operation, peak voltage transients, and switch mechanisms. Simulation of inductive loads is discussed as a means of obtaining more realistic test results.

THE required performance and reliability are not always achieved in present aircraft switch designs. This discrepancy does not result from the difficulty of the problems involved. Given time, equipment environment requirements as well as weight and space limitations can usually be designed to the optimum of refinement. The dynamic, constantly changing nature of the aircraft problem is the principal factor in producing the variance between the qualities desired and qualities achieved. Standards which yesterday were considered impractical or impossible are being met today, only to be replaced by even higher standards for tomorrow. The difference between production design performance and present design objectives may therefore be regarded as principally a phase difference, often unavoidable with such rapidly changing environmental and performance objectives. While the need for higher performance equipment is recognized in the aircraft fields of power generation and power application, it is not always apparent that aircraft circuit interruption and closure equipment is regarded with the same degree of seriousness. This condition, of course, will correct itself in due time, but it is suggested that at present the industry may be in a period where this phase lag between production switch gear and the anticipated requirements of aircraft design is likely to be especially large and critical. Therefore, a discussion of switch evaluation and testing techniques may be timely.

Evolutionary changes are to be expected. In addition to changing environmental requirements of lower pressures and higher temperatures, aircraft switch design and testing procedures are likely to be strongly affected by:

1. A critical need for greater reliability.
2. The circuit interruption transient voltage problem. Damage from peak voltage transients that cannot be solved entirely by improved insulation poses a severe hazard to high-altitude operation and control of inductive aircraft electric equipment. It may not be practical to increase surface insulation creepage distances sufficiently to provide safe extreme-altitude operation for the several-thousand-volt inductive circuit interruption transients now generated by present switch designs.

Reliability Requirements

Increased reliability is needed in aircraft switchgear if it is to keep pace with aircraft advancement. Insulation test requirements are an example. For aircraft electrical switches and other equipment these insulation requirements are considerably below the insulation test requirements of either the electrical industry or the Underwriters Laboratory. The electrical industry and Underwriters Laboratory both require high potential tests in excess of 1,000-volt rms 60-cycle to meet minimum safety requirements even for household equipment. Much air-

---

Paper 54-352, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing July 30, 1954.

T. R. STUELPNAGEL is at 3140 Hollypark Drive, Inglewood, Calif., and J. P. DALLAS is at 8511 Vicksburg Avenue, Los Angeles, Calif.

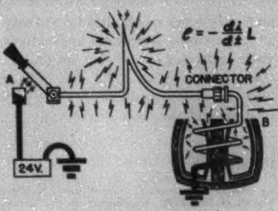


Fig. 1. Voltage stress caused by circuit interruption

craft electrical equipment is tested to only 500 volts rms 60 cycles.

High over-all system reliability requires a very high reliability of the individual components of the system. For example, if there are 100 switches on an airplane, all of which must operate correctly to complete a military mission, then to have a probability of failure of one switch out of 100 flights means that the hypothetical airplane must be equipped with switches that have a life expectancy of 10,000 operations without a malfunction. Testing and inspection procedure to ensure such a degree of reliability represents a considerable change from present practice.

Positive operation contacts are essential in switches where high reliability is needed. Switches that limit the travel of landing gear actuators or gun turrets are specific examples. Failure of such switches to open can result in serious mechanical damage to the airplane. The fire hazard of the resultant stalled motor must also be considered. In aircraft switches the emphasis has been placed on snap-action contact separation, with little effort expended in providing positive mechanical follow-up on these snap-action mechanisms that could, in an emergency, open the contacts with the many pounds of force available in the typical operating device instead of the few ounces of force available in the typical snap-action mechanism. Snap-action contact separation may be desirable from an electrical standpoint; however, mechanical follow-up on the snap action is essential to really reliable switch operation.

Inductive Circuit-Switching Transients

Several -thousand-volt inductive surges may result when a circuit is interrupted to inductive aircraft equipment such as solenoids, relay coils, motors, etc.[1] These voltage transients in excess of 1,000 volts for 28-volt aircraft equipment appear as a severe voltage stress to the insulation of the switch as well as the equipment generating the surge. Fig. 1 is intended to illustrate this condition. Normally the instantaneous breakdown impedance of the switch gap during circuit interruption determines the value of peak voltage generated. However, with aircraft equipment insulated and tested to only a 500-volt rms 60-cycle high potential test, it often happens that the transient voltage damages insulation in the equipment or switch before it can be harmlessly dissipated in the switch gap. Measurements demonstrate that circuit-opening voltage transients are as severe in 28-volt aircraft equipment circuits as in systems of higher voltage with similar equipment.

Circuit interruption time may be a critical factor in aircraft switch application. Consideration of the expression for the electromotive force of self-induction: $e=-(Ldi/dt+idL/dt)$ where e=transient voltage, L=inductance, i=current, and t=time, will show that the potential voltage stress of the switch and equipment insulation is proportional to the instantaneous rate of change of current and inductance.

Therefore it is not enough merely to limit total interruption time to some reasonable figure. For example, it is commonly recognized that when the total circuit interruption time approaches a millisecond or less per ampere of interrupted load, destructive voltage transients will occur. (This is a condition found only in the gross misapplication of switchgear such as the ill-considered use of vacuum, oil, or solid dielectrics in the gap of rapid action switches.) However, a much more common and critical, though seldom recognized, condition is the fact that voltage transients of several thousand volts may occur in a circuit with a total interruption time of many milliseconds per ampere.

These high-voltage transients are caused by discontinuity near the end of the interruption cycle. Fig. 2 is a typical representation of voltage and current conditions during circuit interruption. It is to be noted that the sharp current discontinuity near the end of the interruption cycle produces the principal voltage rise.

This final microsecond discontinuity in a total interruption cycle of several milliseconds is the cause of an insulation-puncturing spike. This is not an exceptional condition. It is the typical result of present switch and circuit design. The factors controlling the value of this transient are many. While a quantitative study has not been made, it can be stated with some certainty that in common aircraft inductive circuits the values of inductance (as it is commonly measured) and resistance are of less importance than the following three factors in determining the peak value of such transients:

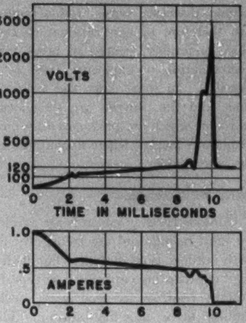


Fig. 2. Typical peak voltage transient

1. Dielectric losses.

2. Core losses. Since such transients are of high frequency and contain little total energy, the peak value is considerably affected by both core and dielectric losses. It has been noted that the substitution of a laminated core for a solid core may more than double the peak voltage measured.

3. The velocity and length of the switch gap during the terminal spike forming part of the interruption cycle.

High-Power Switch Insulation

The insulation of primary power switching gear raises a serious problem where high current faults are possible. Martin and Hauter[2] call attention to the fact that phenolic materials and many other commonly used insulating materials have a tendency to burn vigorously and to contribute, extend, and reinitiate power arcs. Switchgear terminals and contact assemblies connected to main busses and/or high-capacity generator feeders can become involved in this very serious hazard. It is indicated that standard American Society for Testing Materials insulation arc resistance test methods may not be adequate to cover all aircraft conditions. Four-hundred-cycle 115/200-volt aircraft systems of 60 kva capacity or higher are considered in this study. The hazard arises not so much from normal circuit operation as from abnormal or accidental conditions. A misplaced piece of safety wire, loose terminal connections, floating washers, and even tools, inadvertently left near or in junction boxes, have been the cause of serious faults and battle damage. The question is whether, when such a fault occurs, the whole insulation will burn up or the fault clear itself before serious damage results. While investigations of the extent undertaken by Martin and Hauter have not been carried out with d-c systems, preliminary information indicates that almost an identical problem exists where batteries are used to back up a 28-volt aircraft system. And, of course, with 120 volts d-c the situations described in reference 2 can be realized from a normal 30-kw system without battery.

The purpose herein is not to deal with this problem but to analyze it and to suggest that it deserves more attention. The implication of present data is that maximum effort should be made to eliminate phenolic and other arc-supporting materials from heavy power-handling switchgear and to substitute glass-bonded mica and other arc-resisting and nontracking insulators.[3]

Switch Testing

What is the rating for a given switch for aircraft application? The question looks deceptively simple but this impression is far from correct because electrical switching is a complex subject. Perhaps there is no other basic electrical engineering science so little documented, yet troubled with so many variables, as circuit interruption.

In the question of switch testing, the following remarks might be considered representative of types of approach which are not uncommon: 1. "Establish a rating for the switch; test it, but not to destruction, and have the answer by tomorrow." 2. "Extensive data on the performance of switches are unnecessary as long as they meet the specification." 3. "A switch test should be a life test of a few switches with a given set of conditions."

Actually a life test, or a test to the minimum requirements in a specification, constitutes only a small part of proper switch evaluation. Furthermore, the results from a life test alone have limited application unless the basic switch characteristics have also been determined. The tail-end testing used in most switch evaluation has long since been discarded in other fields. It remains in switch testing because a switch is often visualized as being so simple that an organized "scientific" approach is not necessary. Exactly the reverse is true; an aircraft switch test must control and evaluate more than 10 variables. Just as much diligence and engineering science must be used in switch testing as in tests of seemingly more complex equipment. Aircraft switches should be evaluated in the test laboratory to determine these characteristics:

1. Current interrupting capacity at specified voltage, altitude, and type of electrical loads. The range-of-altitude conditions to be considered should be specified because some types of switches may be critical at certain altitudes while operating better at maximum altitudes or minimum pressure.

2. Current-carrying capacity should be determined at the most critical condition that is usually found from a study of maximum ambient temperature and altitude requirements.

3. Peak interruption transient voltage resulting from the interruption of specified aircraft inductive loads should be determined so that, if necessary, provisions may be made to prevent damage to insulation of the switch itself and to equipment on the airplane.

4. Life expectancy of the switch should be determined at rated current and specified operating conditions.

VARIABLE FACTORS AFFECTING SWITCH TESTING

In a typical switch test there are 14 factors or conditions that must be either controlled or measured if the test is to be conclusive. The seven conditions which must be controlled are: current; voltage; type of load; atmosphere in which switch operates; frequency of switch operation; mode of operation, i.e., the speed and pressure with which the operating element of the switch is moved; and type of current—a-c, frequency, d-c, and supply system regulation under the specified load.

The other seven factors are characteristics of the switch. They may be dependent upon the foregoing conditions but should be monitored throughout the tests. They are: arc time; peak voltage during interruption cycle measured across the switch gap with a recording device having an input circuit impedance of at least 1 megohm and a frequency response of 1 megacycle; contact heating; arc energy; speed of contact separation; contact pressure; and contact separating force.

The tendency in switch testing is to neglect some of the variables completely and to discount the importance of others. For example, arc time and its relation to maximum interrupting capacity is not usually evaluated in a switch test. Peak voltage, arc energy, and mode of operation are other variables that may seem unimportant but actually must be known or controlled to make a valid switch test.

THREE PHASES OF SWITCH EVALUATION

The test of a switch may be divided into three phases:

1. Mechanical evaluation of the switch mechanism and the contact assembly.

2. Electrical test of switch to determine current-carrying and interruption capacity as functions of the specified electrical supply, environment conditions, and load.

3. Life test of switch at rated current and specified load.

It is to be noted that conditions for proper switch life tests are dependent upon the results of the foregoing tests.

THE SWITCH TEST REPORT

A complete aircraft switch test report should present the following data:

1. Observations, photographs, and measurements defining the mechanics of the switch.

2. At least two curves of arc time as a function of current.

3. Four curves of current-interrupting capacity as a function of various altitudes verus load.

4. One curve of current-carrying capacity showing the millivolt drop across the closed contacts as a function of current.

5. Results of life tests.

Mechanical Evaluation

Before starting electrical tests on a switch, several units should be disassembled and the mechanics of the switch studied. Some general criteria for evaluating mechanical reliability are:

1. There should be general simplicity and workmanship of the unit.

2. Contact closure and separating forces should be measured and should be adequate. In general, if forces available to open and close contacts are less than 1 ounce they are inadequate. Analysis of these observations with respect to vibration and acceleration should be made to determine if the net contact forces are likely to be adequate under the extremes of operational environment.

3. Contacts should close with a mechanical overtravel of at least 0.010 inch. This preload provides tolerance for contact erosion.

4. The open contacts should provide a gap of at least $^{1}/_{16}$ inch to ensure freedom from malfunctions caused by metal transfer, contact bounce, and mechanical tolerances.

5. Contact points should be securely fastened to contact arms and should have a proper thickness of contact metal to allow for anticipated erosion.

6. Nontracking arc-resistant insulation materials should be used, with special attention to the nature of the material and clearances around the contact gap and terminals, where arcing and spark-over could occur.

7. The melting or softening points of all materials used should be considered. In general, soft solder and plastics that soften at temperatures below 400 degrees Fahrenheit should be questioned.

8. The construction of detent mechanisms should be especially noted. They are the most complicated part of a switch, and con-

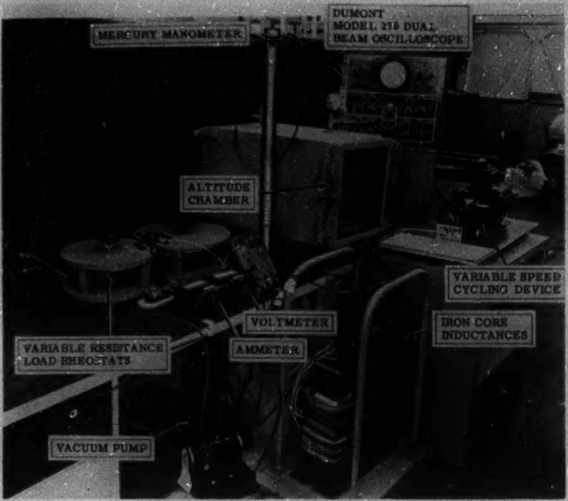

MERCURY MANOMETER
DUMONT MODEL 218 DUAL BEAM OSCILLOSCOPE
ALTITUDE CHAMBER
VARIABLE SPEED CYCLING DEVICE
VOLTMETER
IRON CORE INDUCTANCES
VARIABLE RESISTANCE LOAD RHEOSTATS
AMMETER
VACUUM PUMP

Fig. 3. Test setup for switch analysis

sequently the most likely to fail because of mechanical wear, temperature, or production tolerances.

If there is any doubt about the mechanical ability of the switch to withstand all operating conditions and to operate the required number of cycles, a preliminary accelerated mechanical test should be run before starting the electrical test. Time and money can be saved by eliminating unsatisfactory mechanical designs early in the test schedule.

It is the writer's opinion that aircraft switches should be tested for a mechanical life of not less than 100,000 operations, if really reliable 10,000 to 20,000 cycles of operation are to be ensured for the production product installed in an airplane.

Electrical Evaluation

Fig. 3 illustrates a typical assembly of switch test equipment consisting of a dual-beam oscilloscope, a pressure-controlled chamber, a resistance load bank, variable inductors, a variable-speed cycling motor and a recording camera for the oscilloscope that includes, preferably, a Land polaroid processing unit. The Land camera permits immediate evaluation of recorded transient data. These items of equipment may be arranged in a test circuit as shown in Fig. 4. Fig. 5 illustrates a typical oscillogram resulting from such an assembly. One of the special advantages of this arrangement is that it presents on one oscillogram the variables of current, voltage, arc time, arc energy, peak interruption transient voltage, and load inductance-resistance (L/R) ratios.

Determination of Switch Arc-Time Characteristic

One of the revealing characteristics of a switch is the variation of arc time with current. In general, as current through the switch is increased, the arc time on contact separation will also increase slowly. However, when the current approaches a critical value, the arc time suddenly increases rapidly with a small increase in current until the arc sustains. Once this curve has been established for a specified set of conditions, the switch may be rated quite accurately by picking a current value below the knee of this curve. A ratio of about 75 per cent of the value determined for the knee of this arc-time versus current curve is regarded as a practical, safe rated current for most aircraft switches.

It would be desirable to have arc-time versus current curves for several values of inductive loads and also for increments of altitude; however, the large combination of conditions required to obtain these curves is impractical for the typical switch test. A more realistic procedure is to use only two arc-time curves as rough guides in obtaining the more important altitude characteristic curves to be obtained later. It is suggested that the two arc-time curves be taken at sea level, one curve with a resistive load and the other with an appropriate inductive load.

Determination of Switch Altitude Characteristic

Four curves of maximum switch-interrupting capacity versus altitude, for the four folllowing conditions, are suggested: First, a resistive load curve plotting maximum currents to 50,000-foot altitude may be run. Then, three inductive load

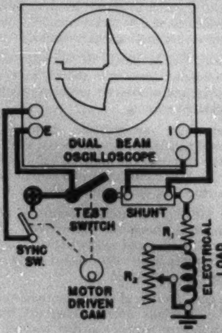


Fig. 4 (left). Test circuit used in switch analysis

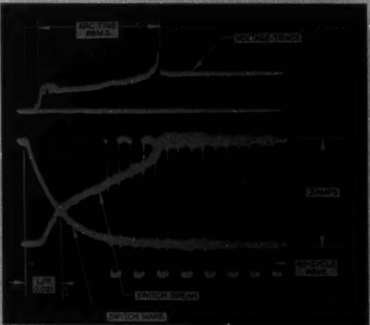


Fig. 5 (right). Oscillogram of switch operation obtained with the test setup of Figs. 3 and 4 curves to 50,000 feet at three different values of L/R, are desirable. Since the objective in a switch test is to obtain a switch current rating under specified conditions of altitude and load, it follows that the major portion of the time allotted for the test should be spent in this effort.

In the resistive load test, while cycling the switch at 25 cycles per minute and observing the arc time on the oscilloscope, the current is increased until the critical arc time is approached. The critical arc time for the sea level condition can be estimated from the arc-time curves previously obtained. In a similar manner, critical current values are obtained for altitudes up to 50,000 feet. It is suggested that 5,000-foot increments of altitude are adequate for rating the switch. Similar curves are then run with loads selected to cover the L/R range of 0.005 to 0.03.

Current-Carrying Capacity

Current-carrying capacity is determined by the I^2R loss in the contacts themselves. This power dissipation can be measured by running a curve of millivolt drop across the closed contacts as a function of current through the contacts. Ordinarily, the current-carrying capacity of the contacts will exceed the current-rupturing values. But should the contact heat generated at rated break current be excessive, the switch will have to be rated on the basis of current-carrying capacity. Acceptable contact temperatures or power dissipation values are dependent upon the materials and type of switch involved. In general, contact current-carrying capacity tests should not be made on new contacts but should be run after 50 per cent, and again, at completion of the life cycle test.

Switch Life Test

Once the rated current for the switch has been established from the altitude versus current-carrying-capacity curves, life tests can be run at boundaries of current, altitude, and L/R ratio established by these data. In setting up life tests, the 14 factors or conditions outlined previously must be given consideration and adequately controlled or recorded. For any test aimed at really evaluating a switch, the minimum life cycle operations required by applicable specifications should be considered, but not regarded as an objective or limit. In general, the fewer samples available to establish satisfactory operation, the larger should be the number of life cycles. This is because, with a large number of cycles, disturbing variables are more likely to be easily determined. But where critical use is to be predicated on very limited test work, the boundaries of these limited test requirements must be set much higher to ensure an equal degree of security.

The number of switches required for a test evaluation will depend upon the completeness desired for the test. Twenty is suggested as a practical number of test switches. Ten of these switches may be used to determine characteristics and the other ten should be life-tested at the conditions established or specified.

Performance Testing Versus Engineering Evaluation

In the discussion of switch testing the premise was initiated that the problem was to find the rating for a switch. This problem is rightfully separate and different from the problem of testing for specification conformity. The proposed test is outlined with no intent to depart from the accepted military specification policy of stating all requirements in terms of required performance. However, following a specification policy does not mean that full-scale engineering tests and consideration of switchgear are unnecessary or undesirable. The designer and applications engineer should know much more about a switch than would be justifiably required in a military specification.

Inductive Load for Switch Testing

Of all the convenient approximations used by test engineers to describe operating conditions, probably none is more abused than the L/R ratio practice of describing inductive switch circuit loads. In fact, the simulation of aircraft circuit inductive loading for test purposes is far from acceptable. First, very little data on the nature of aircraft inductive loads are available and, second, it is doubtful whether existing terminology can be used without some invention to express accurately the nature of such iron-cored inductances where the iron is nearly always operated near saturation.

The L/R ratio does not define an inductive switch load in a useful manner if it is determined from the current rise time or any other method that gives a determination of L on the basis of low frequency changes in the circuit. The principal circuit interruption voltage peak occurs near the end of the interruption cycle, and is a relatively high-frequency phenomenon of 10 to 200 microseconds' duration. The voltage rise rates are of the order of 100 volts per microsecond. For these conditions, low-frequency values of L have little significance. Dielectric losses and core losses are likely to be dominant factors in determining the peak voltage and energy values.

The optimum solution of the inductive load problem is to life-test the switch with the actual load which it is to operate on the airplane. When, as is usually the case, the actual load cannot be used in a switch test for reasons of excessive heating, mechanical limitation, or expense, an electrical mock load is the next best solution. But this mock load, to be equivalent, should be adjusted to simulate actual aircraft equipment circuit interruption transients. In Fig. 4, the load inductance L may be an air-core coil of low distributed capacity and high-voltage layer insulation. The simulated load is adjusted with an oscilloscope to give a current rise time approximating the class of aircraft equipment for which the switch is being tested. But finally and most important, the resistance R shunted across the inductance should be adjusted to give interruption transients similar to those observed or anticipated from properly insulated aircraft equipment under test. An oscilloscope capable of high-frequency response is required, and the input impedance of the oscilloscope connected across the test switch must be 1 megohm or higher.

Recommendation

Performance and reliability of aircraft switches must be improved if circuit interruption equipment is to keep pace with aircraft requirements. As a first step in achieving this goal, it is recommended that the scientific techniques, so effectively applied to the design and testing of other seemingly more complex aircraft equipment, be also applied to the aircraft switch equipment. Scientific studies of this type will result in the establishment of basic design parameters that will eventually improve switch design and application.

References

1. Peak Voltages with D-C Arc Interruption for Aircraft, V. E. Phillips, W. P. Mitchel. *AIEE Transactions (Electrical Engineering)*, vol. 63, Dec. 1944, pp. 944–49.

2. Arcing Performance of Plastic Insulation, T. J. Martin, R. L. Hauter. *Electrical Manufacturing*, New York, N. Y., April 1954.

3. An Environment-Free 120-Volt D-C High-Performance Limit Switch, Thomas R. Stuelpnagel. *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1289–93.

◆

No Discussion

Symposium on Higher Distribution Voltage for Metropolitan Areas*

The Relative Feasibility of 460-Volt or 208-Volt Service in Commercial Buildings

H. G. BARNETT
FELLOW AIEE

R. A. ZIMMERMAN
ASSOCIATE MEMBER AIEE

H. E. LOKAY
ASSOCIATE MEMBER AIEE

SEVERAL utilities have adapted 480-volt multiple-unit secondary-network installations or spot networks to supply high concentrations of power load, such as air conditioning, in commercial buildings since about 1937. The increasing incidence of these power loads in big buildings in commercial areas and their occurrence in institutional buildings, laboratory and office buildings, and shopping centers in urban and suburban areas encouraged the consideration of some voltage higher than 120/208 volts for supplying such loads. The use of 265 volts for fluorescent lighting in industrial plants suggests that attention should be focused on 265/460 volts as a desirable voltage for supplying the loads in new or rewired commercial buildings.

The relative feasibility of 265/460 volts and 120/208 volts for distributing power in commercial buildings depends on many practical, technical, and economic factors. Most of the practical and technical factors have been resolved by years of experience and practice in many industrial and a few commercial applications of the higher voltage and by the development and production of equipment and devices suitable for the proper operation of systems at this voltage.[1–3]

This paper shows, by a comparison of relative costs of 265/460-volt and 120/208-volt systems, how the nature of the load in a building, the size of the building, and the load density in the building all affect the relative feasibility of the two voltages for serving commercial buildings.

Factors Affecting Feasibility

There are many practical, technical, and economic factors affecting the choice of 460 or 208 volts for supplying the loads in commercial buildings. The practical considerations involve wiring practices, requirements of codes and standards, and safety. The principal technical factors are current levels, voltage drop, interrupting duties, and the electrical suitability of the functional devices involved in distributing and using electric power in the building. Economic comparison should consider the installed cost of the service entrance provisions, the wiring or distribution system in the building, and the utilization equipment and its control.

Practical Factors

Properly installed systems can be as free of hazard to people and property at 460 volts as at 208 volts. A few special precautions are required for proper installation of the 460-volt circuits and these precautions will make minor or insignificant economic differences. The National Electrical Code generally makes no distinction among voltages below 600 in the principal circuits and related equipment in building wiring.[4] The most recent edition of the National Electrical Code permits the use of voltages up to 300 volts to ground on circuits supplying permanently installed electric-discharge-lamp fixtures, provided the luminaires are mounted 8 feet from the floor and do not involve manual switch control as an integral part of the fixture. Hence, a grounded 3-phase 4-wire 265/460-volt system can supply directly all of the fluorescent or electric-discharge lighting as well as motors in a building. Codes and standards probably will present no lasting restrictions on the use of 460 volts for commercial building wiring.

Wiring materials, devices, and practices have been well developed for 460-volt systems by the long-time use of this voltage in industrial plants as well as in several commercial buildings in the last 20 years. Maintenance work on 460-volt building circuits may require a higher classification of electrician in some localities than would be required on 208-volt systems. The same wiring techniques are generally suitable for either voltage.

Technical Factors

The principal advantage of the higher voltage in commercial buildings results from less current and less voltage drop in the circuits. Fewer or smaller circuits can be used to transmit the power from the service entrance point to panelboards or other final distribution points. Smaller conductors can be used in many of the branch circuits serving power loads, and some reduction in other branch circuits may be possible.

It is easier to keep voltage drops within customarily accepted limits in 460-volt circuits than in 208-volt circuits. However, the aggregate voltage drop from the service entrance to the 120-volt loads may be more difficult to keep within limits when these loads are supplied through step-down transformers from 460-volt circuits than when they are supplied from 208-volt circuits because of the voltage drop in the step-down transformers.

The interrupting duty on protective devices may be more severe at 460 volts than at 208 volts for two principal reasons: 1. arcs are less likely to interrupt naturally at the higher voltage;[5] 2. minor elements of impedance normally neglected in fault-current calculations are less effective at the higher voltage. For these reasons it may not be possible in every case to realize the theoretical possibility

---

*The first four papers of the Symposium and their discussions will appear in part III of the Transactions, *Power Apparatus and Systems*, volume 73, December 1954.

Paper S-66, recommended by the AIEE Transmission and Distribution and Industrial Power Systems Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing July 13, 1954.

H. G. Barnett, R. A. Zimmerman, and H. E. Lokay are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

Table I. Data Used in Specific Cases Studied

Case Number	System Voltage, Volts	Building Height, Floors	Lighting Load, Volt-Amperes per Foot^2	Outlet Load,* Volt-Amperes per Foot^2	Air-Conditioning Motor Load, Volt-Amperes per Foot^2	Miscellaneous Motor Load, Volt-Amperes per Foot^2	Elevator Demand, Kva	Estimated Diversified Demand,† Kva
1	120/208	1	1.23	1.6	4.25	0.90		223.6
2	120/208	1	1.23	1.6	5.25	1.44		264.2
3	120/208	1	1.23	1.6	7.50	1.88		340.9
4	120/208	1	2.47	1.6	4.25	0.90		268.3
5	120/208	1	2.47	1.6	5.25	1.44		308.9
6	120/208	1	2.47	1.6	7.50	1.88		385.6
7	120/208	1	5.44	1.6	4.25	0.90		373.5
8	120/208	1	5.44	1.6	5.25	1.44		414.1
9	120/208	1	5.44	1.6	7.50	1.88		490.8
10	265/460	1	1.23	1.6	4.25	0.90		234.3
11	265/460	1	1.23	1.6	5.25	1.44		274.9
12	265/460	1	1.23	1.6	7.50	1.88		351.6
13	265/460	1	2.47	1.6	4.25	0.90		280.2
14	265/460	1	2.47	1.6	5.25	1.44		320.8
15	265/460	1	2.47	1.6	7.50	1.88		397.5
16	265/460	1	5.44	1.6	4.25	0.90		386.2
17	265/460	1	5.44	1.6	5.25	1.44		426.8
18	265/460	1	5.44	1.6	7.50	1.88		503.5
19‡	265/460	1	1.23	1.6	4.25	0.90		223.6
20‡	265/460	1	2.47	1.6	5.25	1.44		308.9
21‡	265/460	1	5.44	1.6	7.50	1.88		490.8
22	120/208	12	1.07	1.6	3.81	1.31	175	2,207.0
23	120/208	12	2.14	1.6	4.80	1.45	175	2,895.0
24	120/208	12	4.71	1.6	6.40	1.56	175	4,405.0
25	265/460	12	1.07	1.6	3.81	1.31	175	2,250.0
26	265/460	12	2.14	1.6	4.80	1.45	175	2,969.0
27	265/460	12	4.71	1.6	6.40	1.56	175	4,428.0
28‡	265/460	12	1.07	1.6	3.81	1.31	175	2,207.0
29‡	265/460	12	2.14	1.6	4.80	1.45	175	2,895.0
30‡	265/460	12	4.71	1.6	6.40	1.56	175	4,405.0
31	120/208	36	2.14	1.6	4.10	1.40	988	8,297.0
32	265/460	36	2.14	1.6	4.10	1.40	988	8,297.0
33‡	265/460	36	2.14	1.6	4.10	1.40	988	8,297.0
34	265/460	36	2.14	1.6	4.10	1.40	988	8,297.0
35‡	265/460	36	2.14	1.6	4.10	1.40	988	8,297.0

* Outlet load served at 120/208 volts in all cases.

† Spare capacity in the panelboards has been included in estimated diversified demand.

‡ Lighting load served at 120/208 volts from 265/460-volt system.

of using lower interrupting-capacity protective devices in the higher voltage system.

Utilization equipment is generally available for either voltage and is electrically suitable for the principal loads in most present-day commercial and institutional buildings. There are suitable 3-phase motors for either voltage level for driving most of the power loads. Fluorescent luminaires and other electric-discharge lamps requiring ballasts are available for 120- or 265-volt supply.

Local or incandescent lighting, small fractional-horsepower motors, and plug-in devices usually are available only for 120-volt operation. If 460-volt service to building loads is used, it is necessary to use step-down transformers to serve these loads. The proportion of this type of load varies considerably depending on the purpose of the building and, to a lesser degree, on the size of the building. The proportion of load that must be served at 120 volts from 120/208-volt circuits has a direct bearing on the relative cost of 460- and 208-volt systems.

Economic Factors

Voltage level affects the installed cost of the elements of a building electric system in various ways. Transformer cost is basically proportional to kva and is affected little by the choice of 460 or 208 volts for the building wiring except that, when the size of transformer units is limited at 208 volts by the practical size of protective devices, the use of 460 volts may permit larger sizes with a correspondingly lower cost per kva.

The cost of low-voltage feeder circuits is approximately proportional to the ampere capacity; hence the aggregate cost of these circuits in a building is less for the higher voltage. However, the branch wiring serving the lighting and outlet loads and some of the small auxiliary loads usually can be the smaller practical size, no. 12 or 10 copper, for either voltage. Most of this wiring is required to reach the fixtures and outlets distributed over the building area. Although the number of branch circuits is substantially reduced at the higher voltage, the cost of these circuits is not reduced proportionally. The outlet wiring is essentially the same regardless of the system voltage because the outlet loads must be served at 120 volts.

The cost of protective devices in the building switchboards is roughly proportional to the aggregate ampere demand of the loads in the building; therefore this cost favors the use of the higher voltage. When the higher voltage is used there will be fewer branch-circuit devices in the panelboards, but the devices will be more expensive than those suitable for 120/208 volts.

The cost of wiring devices depends primarily on the number and location of outlets and the manner of installing the branch circuits. Voltage will have little effect on this cost.

Large motors and their controls cost less at 440 volts than at 220 or 208 volts. Smaller 3-phase motors and fractional-horsepower motors cost the same or less at 220 volts than at 440 volts. Fluorescent luminaires cost practically the same regardless of voltage. Fixtures for the commonly used sizes of incandescent lamps are permissible only on 120-volt circuits (less than 150 volts to ground), hence are not economically applicable directly to 265-volt circuits.

Practical and technical factors present no reasonable deterrent to the use of 265/460 volts for supplying building loads. The main consideration is the relative cost of the electric systems for 120/208 and 265/460 volts.

Economic Analysis

In making a comprehensive study of the relative economics, it is necessary to consider all of the factors mentioned. The economic investigation discussed in the following sections was based on various combinations of loads in three sizes of buildings: 1, 12, and 36 floors. The same floor dimensions of 100 feet by 300 feet were assumed for all floors of all buildings to facilitate the study. The load densities of the various load components and other pertinent load information for the different cases studied are given in Table I. Three densities of lighting, air conditioning, and miscellaneous motor loads were assumed for the 1- and 12-floor buildings. The outlet load was assumed to be constant for all cases investigated. Elevator loads were assumed for the 12- and 36-floor buildings. The data obtained using these variables in the 1- and 12-floor buildings indicated that the use of the average load density in the 36-floor building would give representative results.

The components considered in all comparisons were: the utilization devices and their associated control; the branch circuits; the dry-type step-down transformers used to serve the 120-volt loads from the 460-volt system; the lighting panelboards; the low-voltage feeder circuits

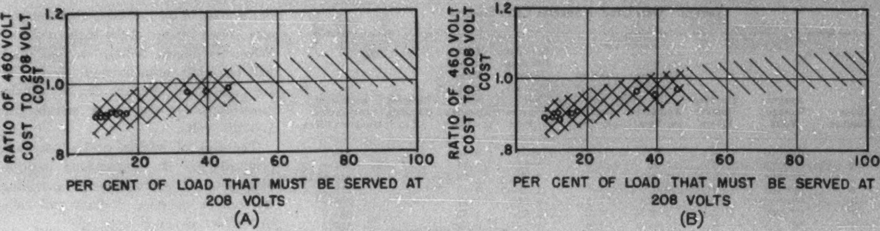


Fig. 1. Relative total system cost for 1-floor building

Double crosshatching includes specific cases studied and the single crosshatched extension is based on judgment and theoretical considerations
A—Primary service. B—Secondary service. This subcaption also applies to Figs. 2, 3, 5, 6, and 7

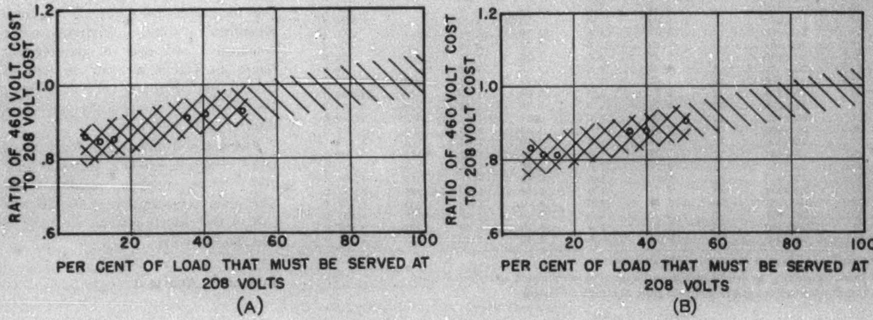


Fig. 2. Relative total system cost for 12-floor building

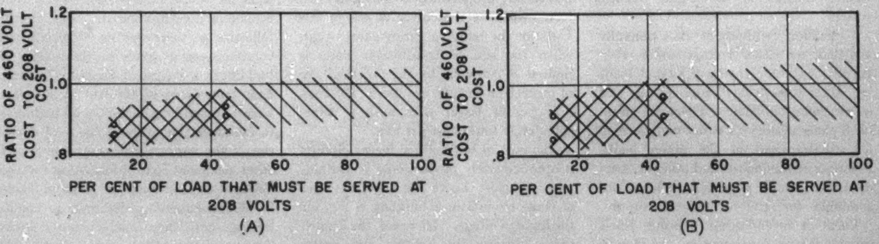


Fig. 3. Relative total system cost for 36-floor building

and bus-duct risers; the building switchboard; and the service entrance which included the service protective devices, the cable or bus run to the transformer units or building switchboard, and the transformer units for the cases where service was assumed to be at primary voltage.

The lighting panelboards and the branch circuits were arranged to provide the proper current-carrying capacities economically and to meet voltage drop limits. The voltage drop in the feeder or riser circuits, from the switchboard bus to the panelboards serving lighting loads, was limited to 1 per cent. The voltage drop in the branch circuits was limited to 2 per cent to the last load. This total feeder and branch circuit voltage drop limit of 3 per cent is both practical and generally accepted.

The miscellaneous motors and air-conditioning motors were assumed in representative numbers and ratings that would be found in practical cases. The costs included in the studies for the elevators were those of the a-c motor and its control. The d-c generator, motor, and other elevator equipment costs were not included.

In the 460-volt cases, step-down transformers were used on each floor to supply the 120-volt loads. The ratings of these transformers were varied consistent with the total load to be supplied at 120 volts in the various cases studied.

Low-impedance bus-duct risers of the necessary capacity, from a voltage as well as a current standpoint, were used in the 12- and 36-floor buildings. Panelboards were supplied by taps from the bus-duct risers at every floor.

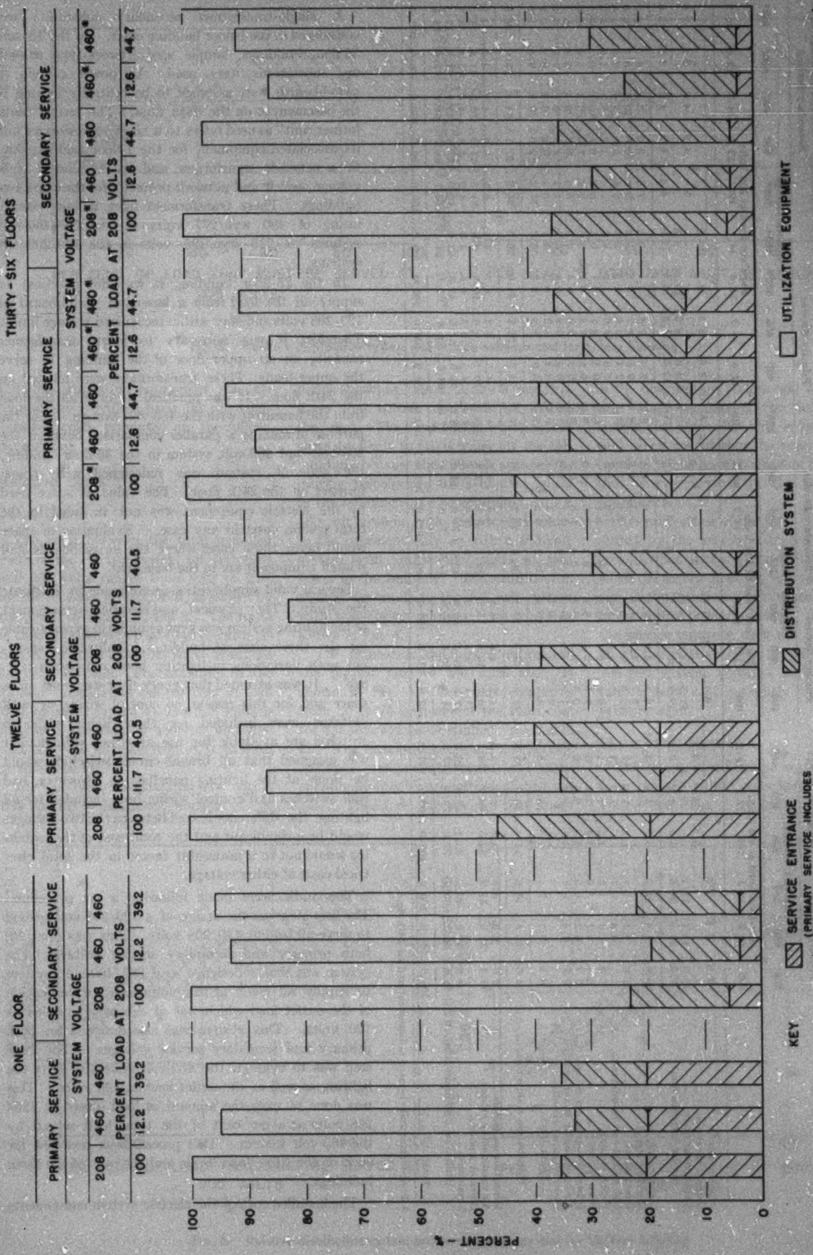


Fig. 4. Comparisons of total system costs and component costs for average load conditions and various building sizes

* TRANSFORMER UNITS LOCATED IN BASEMENT & 24th FLOOR

Table II. Relative Costs of Electrical Components in Commercial Buildings

	1 Floor					
	Primary Service			Secondary Service		
System Voltage	208	460	460	208	460	460
Per-Cent Load at 208 Volts	100	12.2	39.2	100	12.2	39.2
Primary service	7.3	7.7	7.5			
Secondary service				5.5	4.1	3.9
Transformer units	13.3	13.5	13.1			
Service entrance total	20.6	21.2	20.6	5.5	4.1	3.9
Switchboard	5.4	4.4	4.3	6.5	5.4	5.2
Bus-duct risers						
Panelboards	1.0	1.1	1.1	1.2	1.3	1.3
460- to 208-volt transformers		0.8	3.1		1.0	3.8
Feeder circuits	1.9	1.9	1.0	2.3	2.3	1.2
Branch circuits	5.4	4.7	5.6	6.4	5.7	6.7
Miscellaneous motor circuits	0.2	0.2	0.2	0.2	0.3	0.2
Air-conditioning motor circuits	0.7	0.4	0.3	0.9	0.4	0.4
Elevator motor circuits						
Circuit total	8.2	7.2	7.1	9.8	8.7	8.5
Distribution system total	14.6	13.5	15.6	17.5	16.4	18.8
Miscellaneous motors	2.0	2.0	2.0	2.4	2.5	2.4
Air-conditioning motors	33.3	32.3	31.6	39.5	39.2	38.2
Elevator motors						
Luminaires	29.5	31.0	30.2	35.1	37.8	36.7
Utilization equipment total	64.8	65.3	63.8	77.0	79.5	77.3
Total	100.0	100.0	100.0	100.0	100.0	100.0

	12 Floors					
	Primary Service			Secondary Service		
System Voltage	208	460	460	208	460	460
Per-Cent Load at 208 Volts	100	11.7	40.5	100	11.7	40.5
Primary service	6.7	7.8	7.4			
Secondary service				7.5	4.5	4.0
Transformer units	12.3	12.6	11.9			
Service entrance total	19.0	20.4	19.3	7.5	4.5	4.0
Switchboard	8.7	3.7	3.5	9.9	4.5	4.3
Bus-duct risers	6.7	4.3	4.1	7.6	5.1	4.8
Panelboards	2.0	2.6	2.2	2.4	3.1	2.6
460- to 208-volt transformers		1.5	4.5		1.8	5.4
Feeder circuits	0.2	0.1	0.1	0.2	0.1	0.1
Branch circuits	7.7	6.3	8.4	8.9	7.6	10.0
Miscellaneous motor circuits	0.3	0.3	0.3	0.4	0.3	0.3
Air-conditioning motor circuits	1.2	0.9	0.9	1.3	1.1	1.0
Elevator motor circuits	0.1	0.1	0.1	0.1	0.1	0.1
Circuit total	9.5	7.7	9.8	10.9	9.2	11.5
Distribution system total	26.9	19.8	24.1	30.8	23.7	28.6
Miscellaneous motors	4.8	4.6	4.3	5.4	5.5	5.1
Air-conditioning motors	10.9	10.9	10.3	12.4	13.1	12.3
Elevator motors	1.1	1.3	1.2	1.3	1.6	1.5
Luminaires	37.3	43.0	40.8	42.6	51.6	48.5
Utilization equipment total	54.1	59.8	56.6	61.7	71.8	67.4
Total	100.0	100.0	100.0	100.0	100.0	100.0

	36 Floors									
	Primary Service					Secondary Service				
System Voltage	208*	460	460	460*	460*	208*	460	460	460*	460*
Per-Cent Load at 208 Volts	100	12.6	44.7	12.6	44.7	100	12.6	44.7	12.6	44.7
Primary service	4.1	3.3	3.1	3.3	3.1					
Secondary service						4.4	4.2	4.0	2.8	2.6
Transformer units	10.8	9.0	8.5	10.6	9.8					
Service entrance total	14.9	12.3	11.6	13.9	12.9	4.4	4.2	4.0	2.8	2.6
Switchboard	6.7	2.9	2.8	3.1	3.8	7.5	3.2	3.0	3.5	3.3
Bus-duct risers	8.7	9.8	9.2	5.5	5.2	9.8	10.6	10.0	6.2	5.8
Panelboards	2.1	2.4	2.3	2.5	2.3	2.4	2.6	2.5	2.8	2.6
460- to 208-volt transformers		1.6	4.6	1.6	4.7		1.7	5.0	1.8	5.3
Feeder circuits	0.2	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1
Branch circuits	7.9	6.5	8.5	6.7	8.6	8.9	7.4	9.3	7.6	9.9
Miscellaneous motor circuits	0.4	0.3	0.3	0.4	0.3	0.5	0.3	0.4	0.4	0.4
Air-conditioning motor circuits	1.1	0.9	0.8	0.9	0.8	1.2	1.0	0.9	1.0	1.0
Elevator motor circuits	0.4	0.2	0.2	0.2	0.2	0.4	0.2	0.2	0.2	0.2
Circuit total	10.0	8.0	9.9	8.3	10.0	11.2	9.0	10.9	9.3	11.6
Distribution system total	27.5	24.7	28.8	21.0	26.0	30.9	27.1	31.4	23.6	28.6
Miscellaneous motors	4.5	4.5	4.3	4.7	4.4	5.0	5.0	4.7	5.3	4.9
Air-conditioning motors	10.2	10.1	9.6	10.5	9.8	11.5	11.0	10.3	11.8	11.1
Elevator motors	4.4	4.5	4.2	4.6	4.3	4.9	4.9	4.6	5.2	4.9
Luminaires	38.5	43.9	41.5	45.3	42.6	43.3	47.8	45.0	51.3	47.9
Utilization equipment total	57.6	63.0	59.6	65.1	61.1	64.7	68.7	64.6	73.6	68.8
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

* Transformer units located in basement and on 24th floor.

A single-transformer secondary substation was considered in the 1-floor building cases. In the 12- and 36-floor buildings, simple spot-network-type secondary substations were used. All power centers or switchboards were assumed to be centrally located in the basement or on the 24th floor. The term "transformer unit" as used refers to a single transformer and its associated equipment for the 1-floor building, and to a network transformer and its associated high-voltage switch and network protector for the other two buildings. These transformers have a low voltage rating of 480 wye/277 volts in the 265/460-volt systems or 216 wye/125 volts in the 120/208-volt systems.

In the 36-floor building, it was not practical to supply all the load from a basement switchboard at 120/208 volts and stay within recognized voltage limits. Therefore it was necessary to locate transformer capacity on an upper floor of the building to serve the upper loads. These transformers were located on the 24th floor. It was practical to serve all the load from the basement with the 460-volt system. For the purpose of making a parallel comparison between the 208-volt and 460-volt system in the 36-floor building, the 460-volt system was redesigned with transformers on the 24th floor. The value of space used by the electric equipment was not included in the total system costs in any case. Evaluation of space would favor those cases where the switchboards and related equipment are in the basement.

Several valid simplifications were made to facilitate the study. The physical and electrical arrangement of the lighting system was kept as uniform as practicable for the three sizes of buildings for all cases, consistent with variations in lighting level and circuit voltage. It was assumed that every floor was open office space and for this reason no control wiring or wall switches were included for these lights, although switches are available for use at either voltage. It was assumed that all branch-circuit switching would be done at the lighting panelboards. However, had wall switches and control wiring been included for all lighting, the difference in cost between the two voltages would be insignificant and the total cost of the switching would not be a noticeable factor in the total electrical costs at either voltage.

The studies were made following a set procedure. The first step was the design of a 120/208-volt system to serve all load at 120/208 volts. This was done for both primary and secondary service voltages. The system was then redesigned as a 265/460-volt system to supply all loads at that voltage with the exception of the outlet load which was of necessity supplied at 120 volts. This system was also designed for both primary and secondary service voltages. The next step was to redesign the 460-volt system to serve the lighting as well as the outlet load at 120 volts. This was done to vary the amount of load served at 120/208 volts as a per cent of the total load served by the 460-volt system. This process was repeated for each of the three load levels and for each of the three buildings.

The installed cost of the electric system components

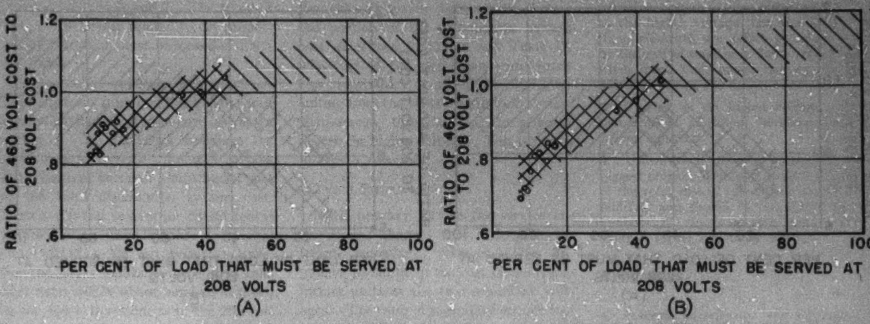


Fig. 5. Relative distribution system and service entrance cost for 1-floor building

and the utilization equipment was estimated to provide a basis for the comparative costs presented in the following sections.

Relative Costs

Total Costs

The ratio between the total installed costs of the electric system for 460 volts and 208 volts ranges from 0.82 to 0.98 for the 35 cases investigated, as shown by the points in Figs. 1, 2, and 3. The trend of this ratio is closely related to the proportion of the total load that must be served at 120/208 volts. For example, in the 12-floor building, as the per cent of load that must be served at 120/208 volts is increased, the ratio of 460-volt and 208-volt cost will increase from 0.81 to 0.92, as shown in Fig. 2.

All the cases investigated are within the double crosshatched portion of the bands shown in Figs. 1, 2, and 3, and it is likely that actual cases will fall within the band. The band within which actual cases are likely to fall is wider for the tall buildings because the 208-volt system, requiring transformer units at two levels in the building, may be compared with a 460-volt system with transformer units in the basement only or with a 460-volt system with transformer units at two levels in the building. The total cost is slightly favorable to the 460-volt system with transformer units at two levels as shown in Fig. 4.

The utilization equipment represents the biggest part of the total cost, as shown in Fig. 4. This part comprises from 50 to 75 per cent of the total cost for the average cases studied; see Table II. The distribution system and service entrance components represent correspondingly smaller proportions of the total cost. Therefore, the savings that result in the distribution system will be significant when considering the distribution system cost alone, but will be less significant when considering the total cost because of the large proportion of total cost that is represented by the utilization equipment.

The proportion of the total cost represented by the service entrance is much greater when primary service is used than when the service is at the secondary voltage, as the transformer units are included in the cost of primary service. However, the service entrance does not materially affect the relative costs of 460-volt and 208-volt system voltages, although the ratio of total cost is slightly higher with primary cost.

Utilization Devices

The saving in utilization cost resulting from the use of 460 volts instead of 208 volts is less than 5 per cent. The large

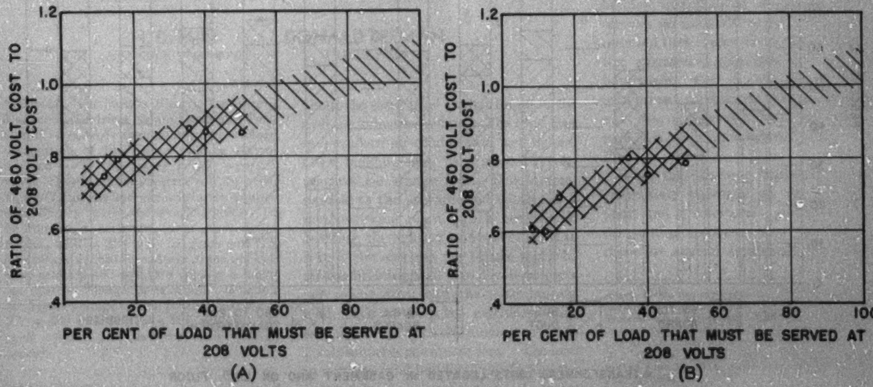


Fig. 6. Relative distribution system and service entrance cost for 12-floor building

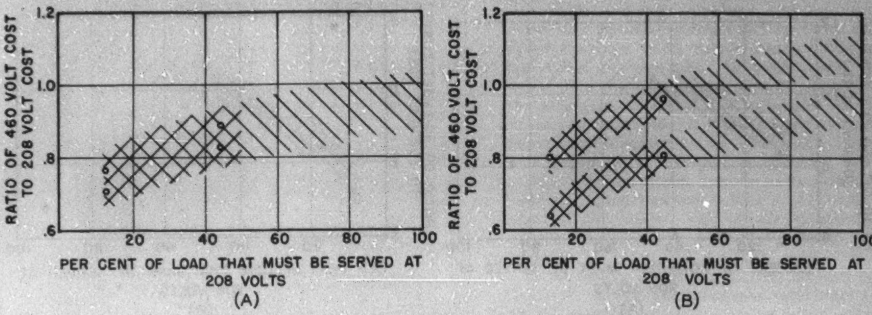
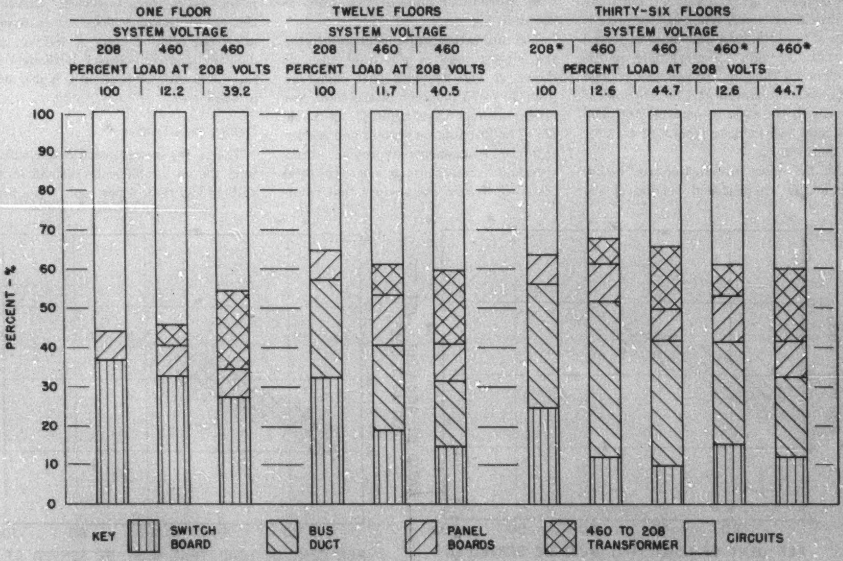


Fig. 7. Relative distribution system and service entrance cost for 36-floor building

motors account for almost all of this small saving, because the cost of the fluorescent luminaires and small motors is the same at both voltages. The luminaire portion is approximately half of the utilization equipment cost for the smaller buildings and approximately 70 per cent for the larger buildings, as shown in Table II. The aggregate cost of miscellaneous motors is also approximately the same for both system voltages because most of these motors are small, but a few large motors, such as drives for fire and water pumps, will be less expensive at the higher voltage. The greater part of the saving found with the 460-volt system in the utilization devices will be in the cost of the larger motors and control used for air conditioning and elevators. Therefore, the saving that results from the cost of the utilization devices at the two voltages will depend upon the amount of large motor load.

Distribution System

The ratio between the cost of the distribution system and service entrance for a 460-volt system to a 208-volt system ranged from 0.59 to 1.04 for the specific cases investigated, as shown by the



*TRANSFORMERS UNITS LOCATED IN BASEMENT AND ON 24th FLOOR

Fig. 8. Division of distribution system cost among the system components for average load conditions and various building sizes

points in Figs. 5, 6, and 7. The trend of this ratio is closely related to the proportion of load that must be served at 120/208 volts.

In the smaller buildings the various circuits represent from one-third to one-half of the distribution system cost, as shown in Fig. 8. In the larger buildings the percentage is decreased because the bus-duct riser cost becomes a substantial part of the total distribution system cost. With a 460-volt system the circuit cost is approximately three-fourths of the circuit cost at a 208-volt system voltage. For the larger buildings the cost of bus-duct risers will be about one-half as much in the 460-volt system as in the 208-volt system except that in tall buildings, where transformer units in the 208-volt system must be at two levels and the transformer units in the 460-volt system can be in the basement only, the cost of bus-duct risers will be about equal.

The difference between the costs of 460-volt and 208-volt switchboards represents the most important saving in the distribution system resulting from the use of the higher voltage. In the larger building the switchboard cost at 460 volts can be as low as 40 per cent of the switchboard cost at the lower voltage. Little or no saving results in the panelboard cost favoring either system voltage.

As the proportion of total load that must be served at 120/208 volts increases, the cost of the step-down transformers in the 460-volt system becomes a substantial portion of the distribution system cost, as shown in Fig. 8 and Table II. The cost of the step-down transformer can absorb part or all of the saving resulting from the other distribution system components, thus favoring to a lesser degree the higher voltage. This is shown in Figs. 5, 6, and 7.

SERVICE ENTRANCE

With primary service the service entrance cost changes very little with different secondary voltages. The difference would lie in the cost of the transformer units at the two secondary voltages. The ratio of secondary service cost in a 460-volt system to that in a 208-volt system will range from 0.5 to 0.85, as shown in Fig. 4. This saving does not have a dominant effect on the total cost.

Conclusions

1. Using a 265/460-volt system to supply the loads in many commercial and institutional buildings is practically, technically, and economically feasible.

2. The total electric system cost is generally favorable to the use of 265/460 volts instead of 120/208 volts when less than about half of the load comprises 120- or 208-volt devices. As the proportion of load requiring 120 or 208 volts increases, there is less likelihood of any reduction in total electric system cost resulting from the use of a 265/460-volt system.

3. The principal elements of the distribution system in the building cost less at the higher voltage, but the savings in these elements are offset in many cases by the cost of small step-down transformers when more than about half of the load must operate at 120 or 208 volts.

4. The use of the higher voltage is more favorable in multifloor buildings than in 1-floor buildings.

5. Utilization equipment accounts for the biggest portion (from half to three-quarters) of the total cost of the electric system in a building and should be considered in any analysis of the economic advantage of using a higher secondary voltage.

References

1. Electric Distribution and Control for Lighting Systems, W. H. Kahler, R. N. Bell. *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 92–99.

2. Distribution Equipment Used on 265/460-Volt Networks and Its Operating Features, L. Brieger, C. P. Xenis, A. J. Bisson, J. DeLellis. *AIEE Special Publication S-66*, Sept. 1954, pp. 20–26.

3. Secondary Network Equipment for 250- to 600-Volt Systems, R. L. Schwab, E. W. Stohr. *AIEE Special Publication S-66*, Sept. 1954, pp. 26–31.

4. 1953 National Electrical Code. *NBFU Pamphlet No. 70*, National Board of Fire Underwriters, New York, N. Y., Nov. 1953.

5. Arcing Fault Currents in Low-Voltage A-C Circuits, C. F. Wagner, L. L. Fountain. *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 166–74.

6. Service Voltage Spread and Its Effect on Utilization Equipment, H. G. Barnett, R. F. Lawrence. *AIEE Special Publication S-66*, Sept. 1954, pp. 51–56.

(Discussions for the Symposium papers appear on pages 334–38.)

Progress in Power System Engineering for Commercial Buildings

H. D. KURT
ASSOCIATE MEMBER AIEE

DONALD BEEMAN
FELLOW AIEE

Synopsis: The Edison Electric Institute—National Electrical Manufacturers Association standard nominal system voltage designation of the higher voltage system referred to in this paper is 480 wye/277 volts. The maximum system voltage is 480 wye/277 and the average utilization voltage will be more nearly 460 wye/265 volts. To be consistent with the lower class designation, such as 208 wye/120 volts which is an average utilization voltage, some utilities have selected 460 wye/265 volts as a nominal system voltage designation. Whether the higher voltage system is called 480 wye/277 or 460 wye/265 volts, it is precisely the same system and has the same voltage spread.

ELECTRICAL loads in commercial buildings have grown in intensity to about the same level as that found in the average industrial plants. The total load in a modern air-conditioned, well-lighted office building may be around 10 volt-amperes per square foot. This is comparable to the average load density in a metal-working plant. In a commercial building, the lighting load will amount to 3 to 5 volt-amperes per square foot, the air-conditioning load 3 to 4 volt-amperes per square foot, and the miscellaneous load, such as business machines, fans, portable lamps, etc., will amount to 1/2 to 2 volt-amperes per square foot. The load in a commercial building may be like the industrial plant, consisting of many integral-horsepower motors, fluorescent lighting for general area illumination, and miscellaneous 120-volt load.

In general, commercial buildings spread out more vertically than does an industrial plant. Even so, a tall commercial building turned on its side electrically resembles an industrial plant. The logical question then is why should the same system engineering principles not apply in both cases. That is, why should small load-center substations located throughout the commercial building not be used just as in an industrial plant. See Fig. 1.

While the foregoing indicates that the generalized approach to an industrial plant and a commercial building power distribution problem may be similar, there are several significant differences.

Paper S-66, recommended by the AIEE Transmission and Distribution and Industrial Power Systems Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 16, 1954; made available for printing July 28, 1954.

H. D. Kurt and Donald Beeman are with the General Electric Company, Schenectady, N. Y.

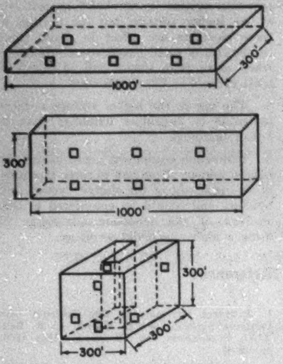

1000'
300'
300'
1000'
300'
300'
300'

Fig. 1. General proportions of factory and office buildings

Some of these are:

1. The extremely high value of floor and volumetric space in a commercial building.
2. The limited space floor-to-floor (usually about 11 to 13 feet).
3. The size of the commercial building both physically and electrically (kva).
4. Many commercial buildings are supplied at utilization voltage from secondary network systems.

While floor space in a manufacturing plant may be as valuable in dollars per square foot of first cost as in a commercial building, the volumetric difference due to higher ceiling heights is significant. In industrial plants, the location of substations is possible over washrooms and in other places where there is no use for the space for production machinery. Such space is not available in commercial buildings. In many cases, the location of substations other than in the basement would require the use of valuable rentable floor space. In some buildings, the value on this floor space has been placed as high as $300 per square foot. This factor will tend to relegate the substation locations to one or two places in the building, and therefore may make it economical to use larger concentrations of substation capacity than are normally found in an industrial plant.

In general, the total load in commercial buildings is of the order of a few thousand kva. The compactness of these buildings usually limits the maximum distance and therefore the maximum length of feeders in the building to less than about 500 feet. There are, of course, exceptions to this, such as the Pentagon Building and the General Office Accounting Building, both near Washington, D.C.; the Merchandise Mart in Chicago, Ill.; the Empire State Building in New York City, etc. However, no such buildings have been built in the postwar period and there are few, if any, so large contemplated. Since the loads are relatively small and the distances relatively short, it is not uneconomical to use a fewer number of larger rated substations than normally are found in industrial plants of comparable kva load.

Table I. Cost Comparison of 20-Story Office Building

	Scheme A	Scheme B	Scheme C
Branch-circuit wiring	$142,400	$161,600	$161,600
Panelboards and transformers	888,300	22,982	140,486
Entrance switchgear and risers	112,200	271,575	112,200
Motor control centers	126,100	174,750	126,100
Four 500-horsepower air-conditioning motors	26,400	29,100	26,400
Motor feeders	5,400	12,510	5,400
Total	$500,800	$672,517	$572,186

Saving of scheme *A* over *B*: $171,717 or $28.6 per kva.
Saving of scheme *C* over *B*: $100,331 or $16.7 per kva.

One of the most significant factors affecting commercial building power system layouts is that many are located in the downtown areas of large cities where power is supplied by the utility only at utilization voltage (i.e., less than 600 volts). As a matter of fact, whether the building is supplied at primary power above 600 volts or supplied at utilization voltage is one of the major differences in considering commercial building power systems. The following discussion will be divided into these two categories.

Whether the buildings are supplied at either secondary or primary voltage, a most important item to be considered from an economic standpoint is the utili-

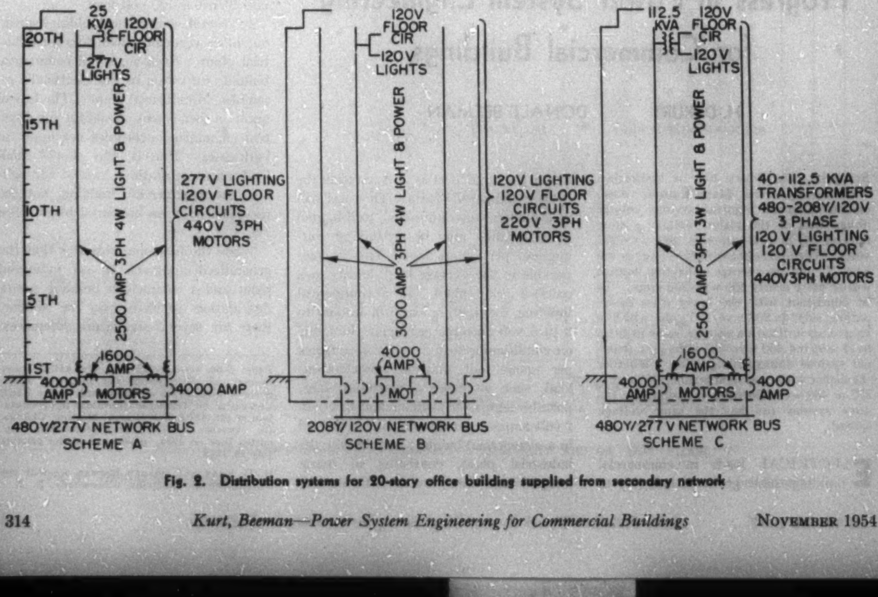

25 KVA
120V FLOOR CIR
277V LIGHTS
20TH
15TH
10TH
5TH
1ST
2500 AMP 3PH 4W LIGHT & POWER
277V LIGHTING
120V FLOOR CIRCUITS
440V 3PH MOTORS
1600 AMP
4000 AMP
MOTORS
4000 AMP
480Y/277V NETWORK BUS
SCHEME A
120V FLOOR CIR
120V LIGHTS
3000 AMP 3PH 4W LIGHT & POWER
120V LIGHTING
120V FLOOR CIRCUITS
220V 3PH MOTORS
4000 AMP
MOT
208Y/120V NETWORK BUS
SCHEME B
112.5 KVA
120V FLOOR CIR
120V LIGHTS
2500 AMP 3PH 3W LIGHT & POWER
40-112.5 KVA TRANSFORMERS
480-208Y/120V
3 PHASE
120 V LIGHTING
120 V FLOOR CIRCUITS
440V3PH MOTORS
1600 AMP
4000 AMP
MOTORS
4000 AMP
480Y/277 V NETWORK BUS
SCHEME C

Fig. 2. Distribution systems for 20-story office building supplied from secondary network

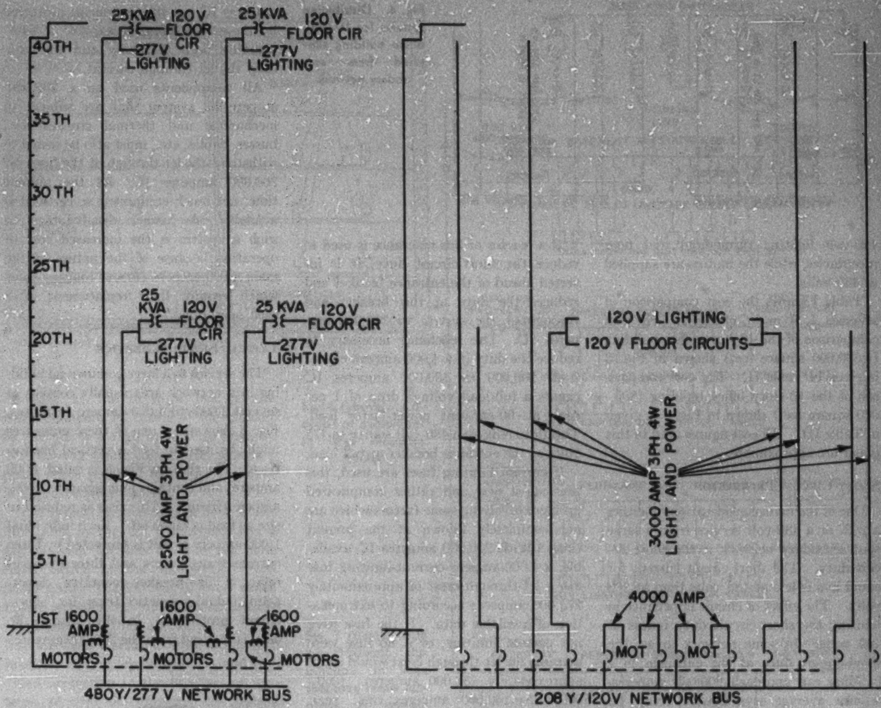

25KVA
120V
FLOOR CIR
277V LIGHTING
25 KVA
120 V
FLOOR CIR
277V
LIGHTING
40TH
35TH
30TH
25TH
20TH
15TH
10TH
5TH
1ST
25 KVA
120V FLOOR CIR
277V LIGHTING
25 KVA
120V
FLOOR CIR
277V
LIGHTING
2500 AMP 3PH 4W LIGHT AND POWER
1600 AMP
1600 AMP
1600 AMP
MOTORS
MOTORS
MOTORS
480Y/277 V NETWORK BUS
120V LIGHTING
120 V FLOOR CIRCUITS
3000 AMP 3PH 4W LIGHT AND POWER
4000 AMP
MOT
MOT
208 Y/120V NETWORK BUS

Fig. 3. Distribution systems for 40-story office building supplied from secondary network

zation voltage. It has been conventional practice to use 208 wye/120 volts almost universally for commercial building power systems regardless of the size of the building. With the development of the 480 wye/277-volt system and 265-volt fluorescent lamp ballasts, it has become economically advantageous in many cases to use the 480 wye/277-volt system. In this system, the standard 440-volt motors are used to drive machinery, and 265-volt fluorescent lamps operated off the line-to-neutral connection are used for general area illumination. Of course, it is necessary to have some power available at 120 volts for supplying the floor outlets and other portable equipment.

Table II. Cost Comparison of 40-Story Office Building

	480 Wye/277	208 Wye/120
Branch-circuit wiring	$ 284,800	$ 323,200
Panelboards and transformers	175,100	45,870
Entrance switchgear and risers	855,310	717,480
Motor control centers	156,900	233,600
Four 1,000-horsepower air-conditioning motors	65,000	80,000
Four 1,000-horsepower motor controls	55,000	60,000
Motor feeders	13,000	30,000
	$1,105,110	$1,490,150

Saving of 480 wye/277 over 208 wye/120: $385,040 or $32.1 per kva.

Commercial Buildings Supplied at Utilization Voltage

COST COMPARISONS

The major part of this discussion on office buildings is whether or not the 480 wye/277-volt system will apply. Economic studies have shown the economics of the higher voltage system in large commercial buildings. The following are summaries of these studies with some extensions.[1-3]

The first cost comparison will be that of a 20-story office building (800,000 square feet) in Fig. 2. Scheme *A* shows the building served from a 480 wye/277-volt network with 277-volt fluorescent lighting and small dry-type transformers supplying the 120-volt floor receptacles load. Scheme *B* shows the building served from a 208 wye/120-volt network with 120- and 208-volt load throughout. Scheme *C* shows the building supplied from a 480 wye/277-volt network with

Table III. Cost Comparison of 10-Story Office Building

	480 Wye/277	208 Wye/120
Branch-circuit wiring	$142,400	$161,600
Panelboards and transformers	88,300	22,982
Entrance switchgear and risers	93,687	218,875
Motor control centers	126,100	174,750
Four 500-horsepower air-conditioning motors	26,400	29,100
Motor feeders	5,400	12,510
	$482,287	$619,817

Saving of 480 wye/277 over 208 wye/120: $137,530 or $22.9 per kva.

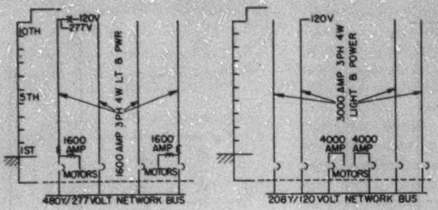


Fig. 4. Distribution systems for 10-story office building supplied from secondary network

120-volt lighting throughout and floor receptacles, while the motors are supplied at 480 volts.

Table I shows the cost comparison of schemes *A*, *B*, and *C* of Fig. 2. The cost comparison of the 40-story office building (1,600,000 square feet) shown in Fig. 3 is given in Table II. The cost comparison of the 10-story office building (800,000 square feet) shown in Fig. 4 is given in Table III. All cost figures used in this paper are approximate.

SHORT-CIRCUIT PROTECTION

One of the main considerations of either a 208- or a 480-volt service from a large spot secondary network is the short-circuit duty. The short-circuit current for equal kva is less at 480 volts than at 208 volts. The effect of circuit impedance in limiting the short-circuit duty is less at 480 volts. In some network areas, the short-circuit duty at the entrance to a building can approach 200,000 amperes, 3-phase average asymmetrical rms. To meet such a requirement, two basic engineering approaches can be followed at 208 or 480 volts, the first being to have all the equipment rated 200,000 amperes interrupting current (IC), and the second being the limiting of the current to 100,000 amperes IC and using existing 100,000-ampere IC equipment.

The use of 200,000-ampere IC equipment presents two problems. The first one is the expensive equipment necessary for such a service; this includes the entrance circuit breakers, the busway, the power panels, and the panelboards throughout the building. The second problem is the unavailability of much of such equipment as standard rated and tested at 200,000 amperes IC.

The limiting of the current to 100,000 amperes IC seems to offer the most advantages. The short-circuit duty can be reduced by the use of reactors or reactance. A second method of controlling the short-circuit stresses is the use of current-limiting, silver-sand fuses that limit the current by their ability to interrupt the current before the crest of the first half-cycle.

If a reactor or bus reactance is used to reduce the short-circuit duty, it is inserted ahead of the entrance breaker and reduces the duty at this breaker and throughout the system to 100,000 amperes IC. The reactance necessary to reduce the duty in a 4,000-ampere circuit from 200,000 to 100,000 amperes IC causes a full-load voltage drop of 1 per cent at 90-per-cent power-factor load. This drop reduces a 480-volt source to 475 volts at the entrance breaker at full load.

If current-limiting fuses are used, this presents a new and rather complicated problem involving some factors which are not completely known at the present time. With 200,000 amperes IC available, a 4,000-ampere current-limiting fuse has a let-through crest of approximately 275,000 amperes according to extrapolation of available data. If the fuse were not current limiting, or if no fuse were present, the let-through crest would equal approximately 452,000 amperes. ($200{,}000/1.25 = 160{,}000$ amperes rms; peak symmetrical $= 1.414 \times 160{,}000 = 226{,}000$ amperes: peak-full offset = direct current + crest alternating current $= 226{,}000 + 226{,}000 = 452{,}000$ amperes.) The crest of the let-through current with a current-limiting fuse is approximately 60 per cent that of the unlimited circuit. The exact effect of the let-through current on many of the components can only be determined at present by test.

There is no doubt in the authors' minds that tests would prove that a 4,000-ampere current-limiting fuse-breaker combination could be designed to withstand 200,000 amperes IC.

The problem associated with the current fuse does not end here. It is carried throughout the entire system. The small fused panelboards with their small current-limiting fuses are subject to a system that has a di/dt of $85.3(10^6)$ amperes per second if 200,000 amperes IC are available. Since the small fuses will attempt to interrupt the circuit much ahead of the larger main fuses, the current-limiting effect of the main fuses may not help reduce the duty at these small fuses. Not only must the fuses be rated for 200,000 amperes IC but the equipment protected by these fuses must be capable of withstanding both thermally and mechanically the let-through current.

All components used on a 200,000-ampere IC system that are subject to mechanical and thermal stresses, i.e., busses, cables, etc., must also be tested to withstand the let-through of the fuses for 200,000 amperes IC. At the present time, not much equipment so certified is available. A further disadvantage to such a system is the increased cost of operation because of the nature of the main 4,000-ampere current-limiting fuses which require their replacement after each operation.

TYPICAL BUILDING SERVICE

The service to a large commercial building in a network area usually consists of several 2,000- to 4,000-ampere entrances. Fig. 5 shows how one of these entrances might be used to feed a vertical busway riser. The riser in Fig. 5 is rated 4,000 amperes and contains four separate 1,000-ampere circuits. This riser is reduced in size as load is taken off. Each individual 1,000-ampere circuit is protected by three current transformers and three overload relays. The breaker contains direct-acting, dual-magnetic trips for short-circuit protection. In this figure, the duty is reduced from the 200,000 amperes

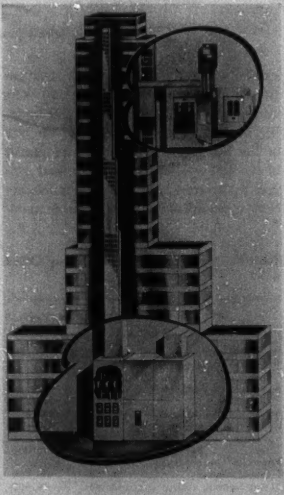


Fig. 5. Typical 480 wye/277-volt distribution supplied from secondary network or from radial transformers

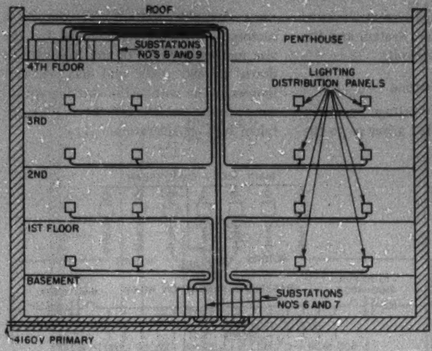

ROOF
PENTHOUSE
SUBSTATIONS NO'S 8 AND 9
4TH FLOOR
LIGHTING DISTRIBUTION PANELS
3RD
2ND
1ST FLOOR
BASEMENT
SUBSTATIONS NO'S 6 AND 7
4160V PRIMARY

Fig. 6 (above). Location of substations in large shopping center

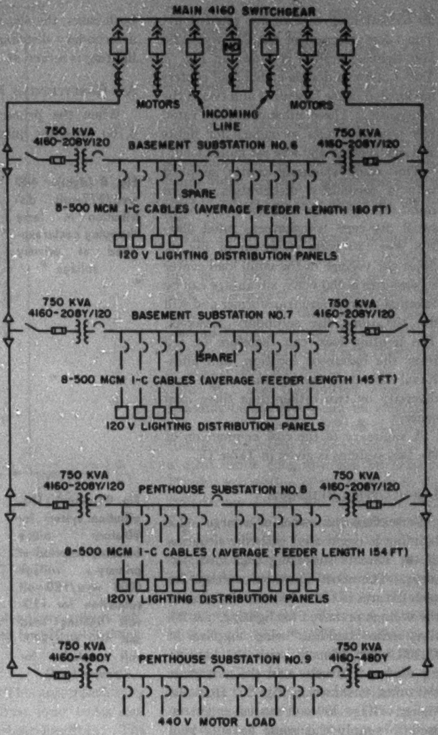

MAIN 4160 SWITCHGEAR
NO
MOTORS
INCOMING LINE
MOTORS
750 KVA 4160-208Y/120
BASEMENT SUBSTATION NO.6
750 KVA 4160-208Y/120
SPARE
8-500 MCM I-C CABLES (AVERAGE FEEDER LENGTH 180 FT)
120 V LIGHTING DISTRIBUTION PANELS
750 KVA 4160-208Y/120
BASEMENT SUBSTATION NO.7
750 KVA 4160-208Y/120
SPARE
8-500 MCM I-C CABLES (AVERAGE FEEDER LENGTH 145 FT)
120 V LIGHTING DISTRIBUTION PANELS
750 KVA 4160-208Y/120
PENTHOUSE SUBSTATION NO.8
750 KVA 4160-208Y/120
8-500 MCM I-C CABLES (AVERAGE FEEDER LENGTH 154 FT)
120V LIGHTING DISTRIBUTION PANELS
750 KVA 4160-480Y
PENTHOUSE SUBSTATION NO.9
750 KVA 4160-480Y
440 V MOTOR LOAD

Fig. 7 (right). 208 wye/120-volt distribution in large shopping center supplied at primary voltage

available at the network bus to the 100,000-ampere IC rating of the breaker by the use of a high impedance busway. This busway would contain the necessary phase space required to reduce the duty in approximately 35 feet.

The typical service to each floor is also illustrated in Fig. 5. This consists of a tap-off from the bus riser to the main fused switch of the 480 wye/277-volt panelboard. The 480 wye/277-volt panelboard is protected by silver-sand fuses rated not more than 150 amperes. This fuse has been tested in combination with breakers having 20-ampere trip and provided protection to the breakers on systems of 100,000 amperes IC. These breakers control the 277-volt lighting circuits, and a fused switch or circuit breaker supplies the small 480–120/240-volt single-phase dry-type transformer, which feeds a 120-volt panelboard to protect the floor receptacle circuits.

AIR-CONDITIONING MOTORS

The air-conditioning motors in large buildings can range from a few hundred to 1,500 horsepower. For motors this size, there is a substantial saving in going from 208 to 480 volts. A further increase of voltage from 480 to 2,400 volts is not generally economically justified. This allows the load in the average building to be optimumly supplied at one voltage namely 480 wye/277.

Table IV. Cost Comparison of Two Systems

	208 Wye/120 Volts, 120-Lights	480 Volts 120-Volt Lights
Lighting substations	$180,891	$101,230
Cable	78,382	23,342
Small dry-type transformers, 480-208 wye/120 volts		73,783
	$259,273	$198,355

Saving with 480 volts: $60,918 or $15 per kva.

Commercial Buildings Supplied at Primary Voltages

When the building is supplied at primary power, there is more freedom in the selection of utilization voltages. For example, in very large buildings such as the U. S. General Accounting Office Building, three voltages have been used: 2,400 volts for very large motors, 480 volts for medium motors, and 120 volts for general area lighting and miscellaneous load. If such buildings were built today, it is very likely that the 120-volt substations would be replaced with 480 wye/277-volt substations.

Some important considerations in designing such buildings are: 1. to select the secondary voltage most appropriate to the load, and 2. to limit the size of transformer to a maximum of about 750 kva at 208 volts and 1,500 kva at 480 volts to keep short-circuit currents within bounds. This may require several substation transformers in one location.

The following are some typical examples of commercial building power systems when supplied at primary voltage.

SHOPPING CENTER

Shopping centers with the majority of the lighting load requiring 120 volts offer a special problem. The shopping center considered here consists of one large main building and many smaller buildings. The smaller buildings are supplied 208 wye/120 volts because their loads do not warrant a higher voltage. The entire area is air-conditioned from one central location. The main building will be considered with 208 wye/120-volt and also 480 wye/277-volt systems. The area most often available for the location of substations is illustrated in Fig. 6. This particular shopping center is supplied at 4,160 volts primary and the large air-conditioning motors are rated 4 kv. The 1-line diagram of the 208 wye/120-volt lighting system is shown in Fig. 7. Two 1,500-kva double-ended load centers are located in the basement and two in the penthouse. Three substations are rated 208 wye/120 volts secondary for lights and one is rated 480 wye for motor load.

This same shopping center supplied with a secondary distribution system rated 480 wye volts with the inclusion of small dry-type transformers rated 480–208 wye/120 volts is shown in Fig. 8. Since the voltage of the main load center secondaries is 480 volts, advantage can be taken of a larger size transformer and still use the same size secondary breakers. The lighting load can therefore be fed from the basement location, and additional advantage can be taken of the load diversity on two rather than three load centers.

A summary of the cost comparison of the two systems is given in Table IV.

Large Office Buildings

Very often the owner of a large office building is confronted with the situation of the tenants supplying their own fixtures. The most universal voltage for such fixtures is 120 volts, and therefore is the voltage preferred for lighting. A 20-story office building being supplied at 13,200 volts primary, with the lighting supplied at 120 volts and the motors at 480 volts, is shown in Fig. 9. How the higher voltage 480-volt system might be used to supply this same lighting load through dry-type transformers is shown in Fig. 10. The voltage drop at the fixtures is approximately the same in each case since the riser drop is less at 480 volts, but the added drop of the transformer must be included. The savings in the 480-volt system is in the riser size and the number of breakers required.

Another important gain in this particular case is space. Fig. 11 shows the floor area required by the two main lighting substations. The only location available for the substation is the basement and the room is broken up by building columns. The 208 wye/120-volt system required a room 120 feet by 20 feet, while the 480-volt system required a room only 70 feet by 20 feet.

The dry-type transformers on each floor are mounted in the wiring closet and occupy the position shown in Fig. 12. This is space that would normally not be used since it is above the recommended height of a manual switch handle.

The savings that can be realized in this building are approximately the same as those found in the shopping center. In both cases, the 480-volt system serves as a secondary distribution system and the lighting is served at 120 volts.

Air-Conditioning Motors

When the primary voltage is above 4,160 volts, the question arises as to the most desirable voltage for the large motors. It is most economical for the size motors involved to step the voltage down to 480 volts and use 440-volt motors. The cost of the starting equipment, feeder cables, and the motors are taken into consideration.

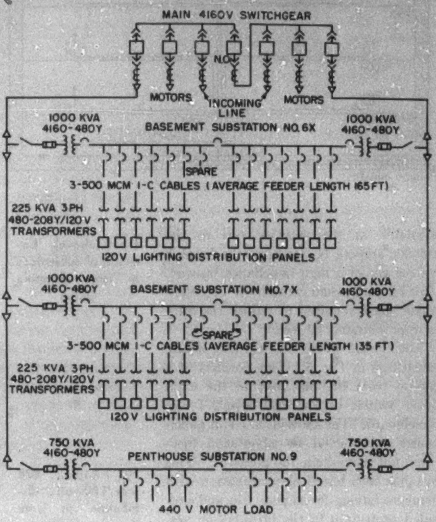

MAIN 4160V SWITCHGEAR
N.O.
MOTORS
INCOMING LINE
MOTORS
1000 KVA 4160-480Y
BASEMENT SUBSTATION NO.6X
1000 KVA 4160-480Y
SPARE
3-500 MCM 1-C CABLES (AVERAGE FEEDER LENGTH 165 FT)
225 KVA 3PH 480-208Y/120V TRANSFORMERS
120V LIGHTING DISTRIBUTION PANELS
1000 KVA 4160-480Y
BASEMENT SUBSTATION NO.7X
1000 KVA 4160-480Y
SPARE
3-500 MCM 1-C CABLES (AVERAGE FEEDER LENGTH 135 FT)
225 KVA 3PH 480-208Y/120V TRANSFORMERS
120V LIGHTING DISTRIBUTION PANELS
750 KVA 4160-480Y
PENTHOUSE SUBSTATION NO.9
750 KVA 4160-480Y
440 V MOTOR LOAD

Fig. 8 (right). 480 wye/277-volt distribution in large shopping center supplied at primary voltage

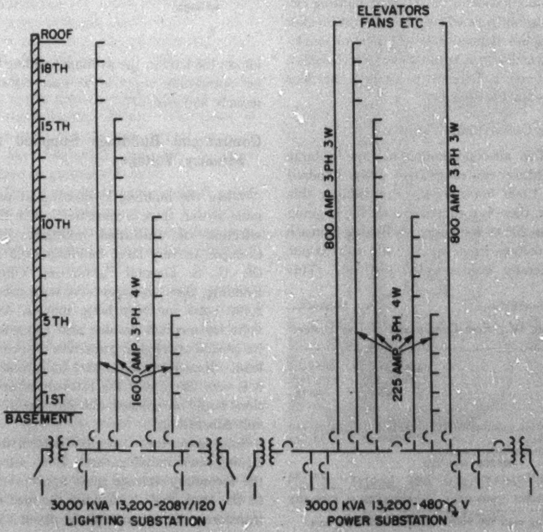

ROOF
18TH
15TH
10TH
5TH
1ST
BASEMENT
1600 AMP 3PH 4W
3000 KVA 13,200-208Y/120 V
LIGHTING SUBSTATION
ELEVATORS
FANS ETC
800 AMP 3PH 3W
225 AMP 3PH 4W
800 AMP 3PH 3W
3000 KVA 13,200-480Y V
POWER SUBSTATION

Fig. 9 (below). Distribution system for 20-story office building supplied at primary voltage (208 wye/120-volt substation for 120-volt lighting load and 480 wye/277-volt substation for power load)

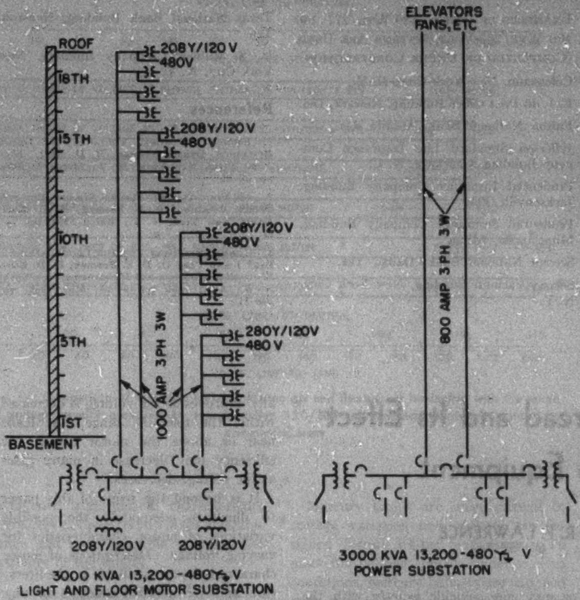

ELEVATORS
FANS, ETC
ROOF
18TH
15TH
10TH
5TH
1ST
BASEMENT
208Y/120V
480V
208Y/120V
480V
208Y/120V
480V
280Y/120V
480 V
1000 AMP 3PH 3W
800 AMP 3PH 3W
208Y/120V
208Y/120V
3000 KVA 13,200-480Y V
LIGHT AND FLOOR MOTOR SUBSTATION
3000 KVA 13,200-480Y V
POWER SUBSTATION

Fig. 10. Distribution system for 20-story office building supplied at primary voltage (all 480-volt substation with 480-volt riser and 480-120/240-volt transformers for 120-volt lighting and 480-volt for power load)

Small Dry-Type Transformers

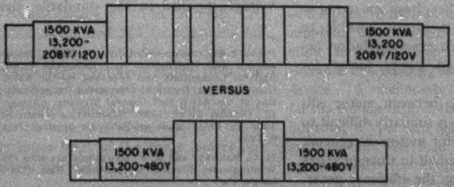
The small transformers included in the 480 wye/277-volt system require some special consideration. The transformers can be of the 2-winding or autotransformer type of construction. Each is reliable, with the autotransformer being less expensive in some instances. With the autotransformer there is a possibility of overvoltages on the secondary service through loss of the neutral or an open winding. The autotransformer is prohibited by some local codes and cannot be used in these localities. The 2-winding transformer might be more expensive, but it serves to isolate the 120-volt circuits and in general is to be recommended.

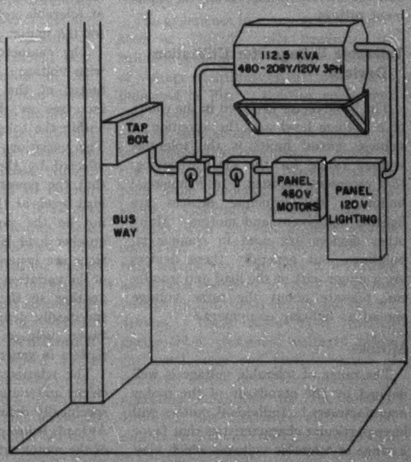
If standard 480-120/240-volt 2-winding transformers are used, they should be equipped with taps below rated primary voltage. Taps are necessary since 480 volts is the maximum for the primary system while 120 is a mean value for the secondary system. A 5-per-cent tap below rated primary voltage will give a ratio of 456 to 120 volts and is the tap to use.

The small transformers, because of their location, require special consideration as to their noise. The noise from such transformers is of two types: the air-borne noise and the transmitted noise. The air-borne noise can be reduced by locating the transformers in a closet and is usually not the source of the major difficulty. The transmitted noise, i.e., sounding-board effect, can be objectionable and yet is not too difficult to control. Resilient mounting of the transformer on a solid inside core wall will be helpful. Flexible connection of the conduit to the transformer is the second requirement. The transmitted noise is important since a very low decibel noise level transformer can be a major source of trouble if improperly mounted. The third consideration for the location of the small transformers is the heat. The over-all losses of the basic 480 wye/277-volt system are less than those of the 208 wye/120-volt system. Provision should be made to ventilate the wiring closets in which these transformers are located. This can be accomplished through the use of louvers with fire dampers that will close with the melting of temperature sensitive link.

Fig. 11 (below). Floor plan of typical 3,000-kva double-ended load centers at 208 wye/120 and 480 wye volts

Fig. 12 (right). Typical wiring closet with 480-volt risers and 120-volt lighting


1500 KVA
13,200-
208Y/120V
1500 KVA
13,200
208Y/120V
VERSUS
1500 KVA
13,200-480Y
1500 KVA
13,200-480Y


112.5 KVA
480-208Y/120V 3PH
TAP
BOX
BUS
WAY
PANEL
480V
MOTORS
PANEL
120 V
LIGHTING

Conclusions

The 480 wye/277-volt system, where applicable, offers many advantages. The savings that can be realized over a comparable 208 wye/120-volt system can exceed $30 per kva of demand. This is due to the reduction in the number and size of breakers used, the reduction in bus size, the lower per-cent voltage drop per unit of current, and the use of more economical motors and starters. The 480 wye/277-volt system can also reduce the space required for a load center due to the use of fewer and smaller feeder breakers for the same load.

EXAMPLES OF WHERE 480 WYE/277- OR 460 WYE/265-VOLT SYSTEMS ARE USED (COMPLETED OR UNDER CONSTRUCTION)

Colosseum, New York City, N. Y.

E. I. du Pont Office Building, Newark, Del.

Fulton National Bank, Atlanta, Ga.

Jefferson Standard Life Insurance Company Building, Charlotte, N. C.

Prudential Insurance Company Building, Jacksonville, Fla.

Prudential Insurance Company Building, Minneapolis, Minn.

Second National Bank, Dallas, Tex.

Socony-Vacuum Building, New York City, N. Y.

Texas National Bank Building, Houston, Tex.

U. S. Rubber Company Building, New York City, N. Y.

References

1. HIGHER NETWORK VOLTAGES IN LARGE BUILDINGS, **Donald Beeman, H. D. Kurt.** *Electrical World*, New York, N. Y., March 8, 1954, pp. 80–85.

2. 480-VOLT SYSTEM MAKES STRONG BID FOR OFFICE BUILDING, **D. L. Beeman, H. D. Kurt.** *Power*, New York, N. Y., March 1954, pp. 74–77.

3. COMMERCIAL HIGH-VOLTAGE INSTALLATIONS—PART I AND PART II, **D. L. Beeman, H. D. Kurt.** *Electrical Construction and Maintenance*, New York, N. Y., April 1954, pp. 94–97, May 1954, pp. 108–11.

Service Voltage Spread and Its Effect on Utilization Equipment

H. G. BARNETT
FELLOW AIEE

R. F. LAWRENCE
MEMBER AIEE

THE adoption of higher secondary voltages in commercial areas[1] involves the careful consideration of suitable voltage spreads and the corresponding plausible voltage levels for supplying utilization equipment. There is urgent need for the industry to select a level that will be generally suitable and acceptable in order to avoid a variety of voltages in various parts of the country and the corresponding variety of equipment ratings.

Tolerable Voltages for Utilization Devices

The basic factor involved in the choice of a voltage level and the adoption of voltage spread limits is the tolerable voltage range for the electric devices that represent the load. In commercial buildings, the principal devices are lighting equipment and motors. Many other devices are used in commercial buildings, but generally these devices are a minor part of the load and usually can tolerate about the same voltage spread as lighting equipment.[2]

MOTORS

The range of tolerable voltage is well defined by the standards of the motor manufacturers.[3] Individual motors will have particular characteristics that favor a range of tolerable voltages which may or may not coincide exactly with the generally accepted range recognized in these standards. However, practically all motors that are likely to occur in commercial buildings are designed to operate at rated output within the standard limits. These limits are 10 per cent below and 10 per cent above the rated voltage of the motor. Motor control devices are also required by standards[4] to meet the same limitations of tolerable voltage as those prescribed for the motor.

The characteristics of a motor vary with voltage. For example, the power factor of the motor demand usually increases as the voltage is decreased within the tolerable voltage limitations. The variation of power factor is also affected by the load on the motor so that the power factor variation at one load level, as the voltage changes, may not be the same as the variation at another load level. While these variations are generally true, the magnitude of the variation differs from one motor to another so that it is not possible to specifically define a precise relation between voltage and power factor for motors in general.

The relationship between motor efficiency and voltage is similarly difficult to specifically define for motors in general. At loads below one-half to three-quarters of the motor rating, the efficiency generally improves as the voltage is decreased within the tolerable range. At loads near or above the motor rating, the efficiency may decrease in many cases as the voltage decreases.

It is beyond the scope of this paper to illustrate adequately the possible variations of motor characteristics for varying voltage. Descriptions of motor characteristics are available in the literature and for detailed analysis of any particular case it is necessary to obtain the characteristics for the particular motor involved to determine just what the variations may be. Experience has indicated that in some industrial operations a voltage at or below the motor rating is preferable to a higher voltage from the standpoint of the efficiency and power factor and the aggregate demand of the motors in the plant. It is likely that most of the motors in a commercial building operate close to their rated output because these motors usually are applied to loads that can be predicted accurately. Hence, there is less likelihood in a commercial building that motors operate below their ratings than in an industrial plant. For this reason it seems desirable to supply motors in commercial buildings with voltages close to or slightly above the rated voltage of the motors.

While motors are designed to deliver rated output within the range of tolerable voltage specified in the standards, the

---

Paper S-66, recommended by the AIEE Transmission and Distribution and Industrial Power Systems Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing July 9, 1954.

H. G. BARNETT and R. F. LAWRENCE are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

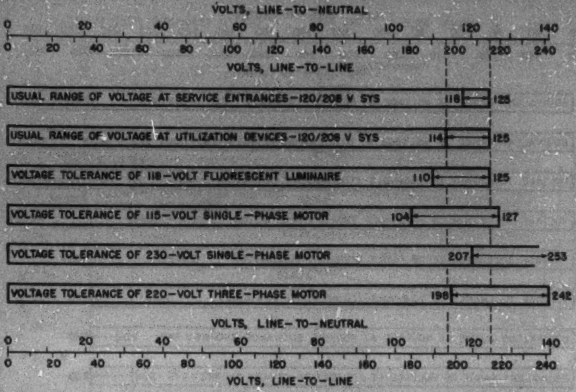


Fig. 1. Comparison of tolerable voltages for motors and fluorescent luminaires with the range of voltage usually provided by representative 120/208-volt secondary-network systems in commercial areas

divergence from normal characteristics is greatest at the extremes of the tolerable range and it is desirable to avoid the extremes of the ranges.

Lighting Equipment[5–7]

The range of tolerable voltage for lighting devices is not standardized by any industry organization, but limits and ratings are fairly well established by the manufacturers of lighting equipment.

The output and life of a filament lamp varies very rapidly with changing voltage. Line voltage higher than rating will shorten lamp life, but will increase light output. A 5-per-cent increase in voltage will reduce life about 50 per cent but will increase light output about 18 per cent. A 5-per-cent reduction in voltage will about double the life but will reduce light output 18 per cent. Therefore filament lamps should be selected with rating nearest to actual line voltage. Filament lamps are available in ratings of 115, 120, and 125 volts.

Fluorescent lamps are not as critical from a light output and life standpoint when operated within tolerable limits. Tolerable limits are approximately ±6 per cent, which has led to the adoption of ballast ranges of 110 to 125 volts, 220 to 250 volts, and 250 to 280 volts for design points of 118, 236, and 265 volts respectively. Voltages higher than the upper limit of range overheat the ballast and reduce ballast life. Lamp life may also be reduced. Line voltage below the lower limit of the range causes uncertain starting.

Mercury lamps are more critical to voltage variation and ballasts are designed for a ±4-per-cent range. To meet the voltage spreads under various conditions, mercury ballasts are supplied with taps that allow the ballast to be operated as close as possible to a ±4-per-cent range.

Voltage Spreads in 120/208-Volt Systems

The present practice in the 120/208-volt systems operated by utilities in this country is a good measure of the desirable spread that should be developed in systems supplying a higher secondary voltage. Suitable voltage spreads at higher secondary voltages should bear about the same relationship to the tolerable voltage ranges of the corresponding equipment as that in the present lower voltage systems. Most utilities maintain voltages within the favorable zone shown in Fig. 1. These voltage conditions generally prevail in normal operation of these systems. It is not practical to keep voltages within these limits during emergency conditions. The preferred voltage report[3] lists the limits at the point of utilization and the system supplying these voltages must necessarily allow for some drop between service entrance to a building and the utilization device. A commonly accepted value for the voltage drop in the building wiring is a total of 3 per cent between the service entrance and the load. This accounts approximately for the difference between the usual range of voltage at service entrances and the usual range of voltage at the utilization device.

Lighting equipment has been designed over a long period of experience to fit very well the usual range of values provided by 120/208-volt systems. The lowest voltage supplied to utilization devices is substantially above the lower limit of the tolerable voltages for fluorescent luminaires. The upper limit of the voltage supplied to the loads is fairly close to the upper limit of the tolerable voltage for the luminaires. This makes the principal operation of lighting equipment at or near the rated voltage of the device. In full-load periods when the majority of luminaires are in operation, the voltage will be in the lower half of the usual range of voltage supplied, but higher voltages in the range generally occur only at light-load periods when only a few luminaires are in operation.

A similar favorable relationship exists between the usual range of voltage supplied by 120/208-volt systems and the range of tolerable voltages for 115-volt motors. Single-phase motors rated at 230 volts do not match the voltages supplied by most 120/208-volt systems because only the upper voltages of the usual range of supply are within the range of tolerable voltages for these motors.

Three-phase 220-volt motors have a range of tolerable voltages that brackets the usual voltages supplied by 120/208-volt systems, but the voltage supplied to these motors will almost always be in the lower half of the range of tolerable voltages for these motors.

The experience in many systems operated at 120/208 volts confirms the suitability of the 114- to 125-volt range of voltage at utilization devices. As is indicated in Fig. 1, lighting equipment and 115-volt motor-driven devices operate entirely satisfactorily. This appears to be a reasonable conclusion in spite of the fact that in a few places there has been some tendency to use 125-volt lamps. Generally speaking, the experience with 220-volt 3-phase motors connected to 120/208-volt supply has been satisfactory. In only a few cases has it been expedient to boost the service voltage for individual motors and related devices. These cases have resulted partly from occasional voltages below the usual range and partly because of marginal motor designs or unfavorable operating conditions for the motors. However, this experience with 220-volt motors on 208-volt systems implies that a slightly higher range of voltage would be desirable.

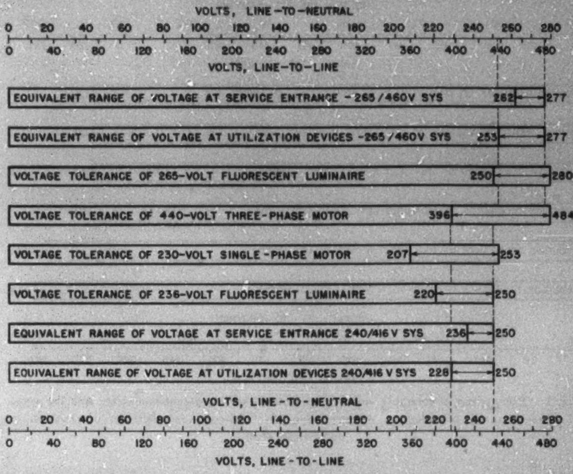


Fig. 2. Comparisons of tolerable voltage for motors and fluorescent luminaires and voltage ranges that can be provided by 265/460- and 240/416-volt systems in commercial areas

Voltage Spreads at Higher Secondary Voltages

The ranges of tolerable voltages for utilization equipment suitable for operation at a higher secondary voltage are illustrated in Fig. 2 and Table I. Ranges are shown in Fig. 2 for two voltage ratings for fluorescent luminaires. Both ratings are available and have been used. The ranges are shown for 440-volt 3-phase motors and for 230-volt single-phase motors.

One of the higher secondary voltages that have been suggested for general use is 240/416 volts, 3 phase, 4 wire, wye.[8] It has been suggested that this voltage would normally operate at about the same relative level as the 120/208-volt systems. A range of voltage at utilization devices giving exactly the same relative voltages as are now considered suitable in 120/208-volt systems would provide a range from 228/395 volts to 250/433 volts. Such voltages could be provided either with transformers having 240-volt windings operating, at no load, at voltages approximately 5 per cent above rated voltage, or by transformers with 250-volt windings corresponding to the use of 216 wye/125-volt transformers in many 120/208-volt secondary-network systems. This range of voltages at utilization devices would fit very well 236-volt luminaires and 230-volt single-phase motors. However, this voltage range is subject to the same hazard when 440-volt motors are supplied as now exists in the use of 208 volts to supply 220-volt motors. Presumably, there should be slightly better operation at the higher voltage because voltage drops in the various circuits would be easier to keep within acceptable limits. The voltage range from 395 to 433 volts seems undesirable for 440-volt motors because efficiency of these motors probably is poorer when operated at voltages in the lower part of the tolerable range in commercial buildings where the loads are likely to be near the motor rating.

Table I. Ranges of Voltages at Different Secondary-Voltage Levels and Tolerable Voltages of Utilization Equipment

System or Device	Minimum Limit Voltage	Nominal or Rated Voltage	Maximum Limit Voltage
120/208-volt system (favorable zone)			
At service	118/204	120/208	125/217
At utilization device	114/197	120/208	125/217
240/416-volt system (favorable zone)*			
At service	236/409	240/416	250/433
At utilization device	228/395	240/416	250/433
265/460-volt system (favorable zone)*			
At service	262/453	265/460	277/480
At utilization device	253/438	265/460	277/480
Fluorescent luminaire			
118-volt rating	110	118	125
236-volt rating	220	236	250
265-volt rating	250	265	280
Motors†			
115-volt rating (1-phase)	104	115	127
230-volt rating (1-phase)	207	230	253
440-volt rating (3-phase)	396	440	484
220-volt rating (3-phase)	198	220	242
Heating appliances‡			
118-volt rating	107 to 110	118	122 to 125
236-volt rating	220	236	248

* Utilization voltage range established by choosing the reasonable maximum voltage and selecting the lower limit equivalent to the lower limit of the 120/208-volt system.

† Rating ±10 per cent.

‡ Representative values.[6]

The other higher secondary-voltage level that has been suggested would result from the use of 480-volt wye-connected transformers supplying a 3-phase 4-wire system. If the suggested maximum voltage limit indicated in the preferred voltage report for 480-volt systems is used as the top limit for the range of voltages supplied to utilization devices, and the lower limit is set proportionally the same as the lower limit in the 120/208-volt system, the resulting range would be from 253/438 to 277/480 volts. This range of voltages at utilization points is suitable for 265-volt fluorescent luminaires and is in the upper two-thirds of the tolerable range for 440-volt motors. It appears likely that this relationship between the range of voltage at utilization devices and the tolerable voltages for the 440-volt motors would be preferable to the relationship corresponding to a 240/416-volt system because of the nature of the loads on motors in commercial buildings. The proposed range of voltage for the 265/460-volt system is not suitable for supplying 230-volt single-phase motors.

Attainment of Voltage Spreads

The voltage drops in the various elements of a supply system beyond the

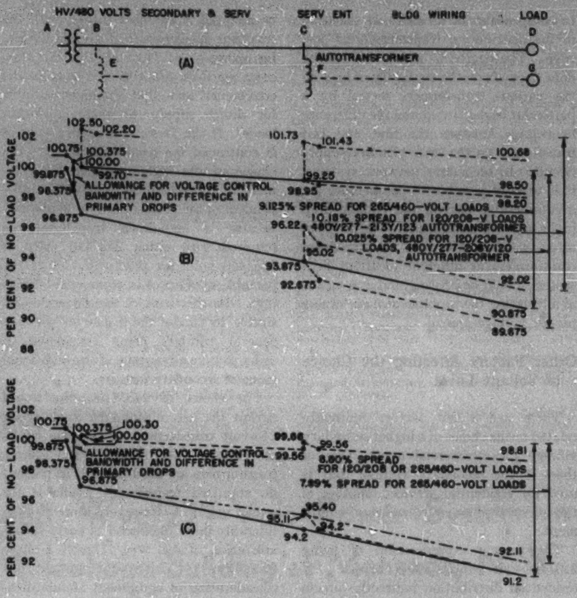


Fig. 3. Allocation of voltage drop in representative system

A. Representative system diagram
B. Voltage profile with conventional autotransformer ratios and with 3-per-cent drop in secondary and service and in building wiring. Drop at light load is one-fourth the drop at maximum load
C. Voltage profile with autotransformer ratio and voltage drops selected to keep voltage spread within favorable limits. Drop at light load is one-fourth the drop at maximum load

point where voltage can be controlled within close limits must be included in the spread between the maximum and minimum voltages at utilization points within the same system. On the basis that the voltage at the primary terminals of transformers supplying a secondary voltage system can be held at a specified level, the spread of voltages at the utilization points in per cent of the no-load voltage is the sum of the per-cent voltage drops in the elements of the system between the primary of the transformer stepping down to secondary voltage and the utilization devices. The elements in which these voltage drops occur then are the distribution or network transformer, the secondary main, the service, the wiring within the building, and any step-down transformer required to change the secondary voltage level. Actually, this extreme limit of the spread does not occur principally because of two things: 1. it is often possible to compensate the voltage control to overcome a part of the drop in the distribution transformer, and 2. the range of voltage drops in the circuits is not from zero to the maximum but from some value corresponding to the light-load level in all of the circuits to the maximum which occurs at full load on the system.

In a distribution system having only one level of secondary voltage such as in the ordinary 120/208-volt secondary-network system, there is no step-down transformer drop to account for. When two secondary-voltage levels are required, as would be the case when some voltage like 265/460 volts is used to supply all of the load in a commercial area although part of the load must operate at 120/208 volts, a step-down transformation would be required. The voltage drop in this step-down transformer then must be included in those drops that must come within the voltage spread at the utilization points.

There are generally accepted arbitrary rules for the voltage drops in certain of the elements in a secondary system. A common value for the drop in the secondary main and service combination is 3 per cent. Likewise, 3 per cent is often used as the limit for the voltage drops in the wiring within a building. If both of these values of voltage drop are accepted, it is difficult to hold the voltages at the utilization points within the favorable zone when all the drops are taken into account. On the same basis, if the voltage drop in a step-down transformer between two secondary-voltage levels must be included, the difficulty of maintaining a favorable voltage spread is further complicated. See Fig. 3 and Table II.

The voltage profile diagram in Fig. 3(B) illustrates the resulting voltage spread when the commonly accepted limits for voltage drops in parts of the system are put together with those in the distribution transformer and a step-down transformer when one is used. Because of the difficulty of maintaining a reasonable voltage spread, it is imperative that the ratio of the step-down transformer be chosen carefully. The ratio must be such that the maximum voltage at light-load periods is at the top of the favorable zone in the lower level of secondary voltage as well as in the higher level of secondary voltage. However, this does not avoid the burden of the drop in the step-down transformer in the problem of maintaining the proper voltage spread in the lower level of secondary voltage.

When voltage drop limits are allocated to various parts of the secondary system on the basis of what can be tolerated within the favorable zone of utilization voltages, it is necessary to limit some of the drops below commonly accepted limits and take advantage of every practicable expedient. This is illustrated by the allocation of voltage drops shown in Fig. 3(C). The voltage profiles in this diagram are based on including the voltage drops in the various elements of the secondary system within the spreads of 114/197 to 125/217 volts or 253/438 to 277/480 volts in the two corresponding levels. For 460-volt loads, the problem is not particularly difficult and results in only a small reduction in the aggregate voltage drops in the secondary mains and building wiring. The illustration is drawn with 3 per cent allocated to the building wiring and the required reduction all made in the secondary mains. When this is done, the permissible voltage drop in the secondary mains up to the service entrance points is limited to 2.7

per cent instead of 3 per cent. To avoid a greater reduction of permissible voltage drop, it has been assumed in this illustration that one-half of the voltage drop in the distribution or network transformer can be offset by the operation of the voltage control devices in the supply system.

For the loads in the lower secondary-voltage level where it is necessary to include the voltage drop in the step-down transformer, the additional drop in this transformer necessitates a further reduction of the aggregate drop in the secondary mains and building wiring. For convenience in developing the illustration, the reduction has been made in the secondary mains only. To minimize the reduction of the drops for the 120-volt loads, a hypothetical ratio was chosen for the step-down transformer so that the ratio of this transformer would give a secondary voltage at no load which would just compensate for the minimum voltage drop in the step-down transformer at light load. Specifically, the ratio in this example is 480 wye/277 to 217.2 wye/125.4 volts where the 100-per-cent voltages at the two levels are considered to be 277/480 volts and 125/217 volts. This hypothetical ratio corresponds approximately to a ratio of 460 wye/265 to 208 wye/120 volts. This case also assumes compensation in the voltage control equipment to offset one-half of the voltage drop in the distribution or network transformer stepping down to 480 volts. In both of these examples, Figs. 3(B) and 3(C), the voltage drop at light load is considered to be one-fourth of the voltage drop at maximum load in every element of the system.

The cases illustrated in Fig. 3 are arranged to illustrate the problem rather than to define the solution of it. It is important that the various factors be carefully considered when the permissible voltage drops for the various elements of the system are allocated. All of the alleviating factors that may be helpful in resolving the problem should be taken into account. If two secondary voltages are to be generally used, it is important that the voltage drop in the step-down transformers between the two voltages be kept as low as practicable. The voltage ratio of the step-down transformer must be carefully chosen to avoid offsetting the voltage spreads in the two levels. It appears desirable to reconsider co-operatively the popular limits for voltage drop in the building wiring as well as in the associated distribution system. Wherever it is practicable, it is desirable to control the supply voltage to compensate for as much of the drop in the distribution transformers as possible. The degree to which this can be done depends on the similarity of load on the various transformers served by a particular feeder circuit and the difference in voltage between the first and last transformer on the same primary supply circuit. In secondary network systems, it is usually possible to compensate for a considerable part of the voltage drop in the network transformer. The prospect of using different taps in the step-down transformers at different locations in a system should be investigated as a means of alleviating the problem of maintaining proper voltage spread.

Other Factors Affecting the Choice of Voltage Level

There are a few factors indirectly related to the choice of a higher secondary voltage for commercial areas. One of these is the secondary voltage that is used for residential service; another is the question of standardization of equipment.

The present consideration of using 240/480-volt single-phase service for residential distribution indirectly affects the consideration of a higher secondary voltage for commercial areas. If a higher residential service voltage is to be adapted in the near future, with the resulting development of plug-in devices for operation at 240 volts, there is then more merit in adapting a voltage in the commercial area that would be suitable for direct supply to these appliance loads. If the present secondary voltage is continued for residential service, the choice of the higher voltage for commercial areas can be practically independent of the voltage used for appliances because a step-down transformer is required with either 240/416 volts or 265/460 volts for practically all of the portable appliances in commercial buildings. In this case, it would seem desirable to choose the higher of the two general voltages being considered in order to take advantage of the additional gains in secondary circuits.

The choice of 240/416 volts would permit the use of standard single-phase 240-volt transformers in 3-phase banks to supply this voltage. If 3-phase transformers are used, there is at present no standard transformer available for either voltage. However, there is considerably more likelihood of early standardization of 480 wye/277-volt 3-phase transformers. This would indicate that consideration of equipment standardization would favor the choice of a 265/460-volt level because 3-phase transformers probably will be used in most installa-

Table II. Voltages at Various Points in Fig. 3. Combination 265/460-Volt and 120/208-Volt System

	Fig. 3(B)		Fig. 3(C)	
Case Units	Per Cent	Volts	Per Cent	Volts
High voltage (*A*)				
Maximum	100.750		100.750	
Minimum	99.875		99.875	
Transformer drop				
Light load	0.75		0.75	
Maximum load	3.00		3.00	
Secondary voltage (*B*)				
Maximum at light load	100.00	480.00	100.00	480.00
Minimum at maximum load	96.875	465.00	96.875	465.00
Service voltage (*C*)				
Maximum at light load	99.25	476.40	99.56*	477.89*
Minimum at maximum load	93.875	450.60	95.11 to 94.2*	456.53 to 452.16*
Voltage at 460-volt load				
Maximum at light load (*B*)	100.00	480.00	100.00*	480.00*
Maximum at light load (*C*)	99.25	476.40	99.56*	477.89*
Minimum at maximum load (*D*)	90.875	436.20	91.93 to 91.2*	441.26 to 437.76*
Voltage at 208-volt loads				
Autotransformer ratio: 480/277–208/120				
Maximum at light load (*E*)	99.70	119.64		
Maximum at light load (*F*)	98.95	118.74		
Minimum at maximum load (*G*)	89.675	107.61		
Autotransformer ratio: 480/277–213/123				
Maximum at light load (*E*)	102.20	122.64		
Maximum at light load (*F*)	101.43	121.72		
Minimum at maximum load (*G*)	92.02	110.42		
Autotransformer ratio: 480/277–217.2/125.4				
Maximum at light load (*E*)			100.00	125.00
Maximum at light load (*F*)			99.56	124.45
Minimum at maximum load (*G*)			91.20	114.00

* The lower minimum values are based on allocating the maximum permissible voltage drops when only 460-volt loads are served. The higher minimum values are based on allocating the maximum permissible voltage drops when both 460- and 208-volt loads must be served. The maximum values are based on serving both 460- and 208-volt loads.

tions. Another factor that has some bearing on the choice is the considerable experience in industrial operations with the use of the 3-phase 4-wire voltage at the 265/460-volt level.

Summary

1. Experience shows that the voltage spread of 114 to 125 volts in 120/208-volt systems has a satisfactory relationship to the tolerable voltage of utilization equipment.

2. Corresponding ranges of tolerable voltages for utilization equipment show that 253/438 to 277/480 volts and 228/395 to 250/433 volts are suitable spreads for 265/460- and 240/416-volt secondary voltages. The spread for 265/460 volts is favorable to 440-volt 3-phase motors and is suitable to 265-volt fluorescent luminaires. Considerable experience has developed with this voltage in industrial plants. An equivalent spread having an upper limit slightly below 480 volts would be more favorable. The higher voltage uses circuit capacity more effectively. The spread for 240/416 volts is favorable to 230-volt single-phase motors and is suitable to 236-volt fluorescent luminaires and appliances.

3. Obtaining a suitable spread of utilization voltage involves a careful allocation of voltage drop limits in the various elements of the secondary-voltage system.

References

1. A NEW APPROACH TO THE PROBLEM OF HIGHER DISTRIBUTION VOLTAGES, A. M. deBellis, S. B. Griscom. *AIEE Special Publication S-66*, Sept. 1954, pp. 3–6.

2. PREFERRED VOLTAGE RATINGS FOR A-C SYSTEMS AND EQUIPMENT. *EEI Publication No. R-6, NEMA Publication No. 117*, Edison Electric Institute: National Electrical Manufacturers Association, New York, N. Y., May 1949.

3. STANDARDS FOR MOTORS AND GENERATORS. *NEMA Publication No. MG 1–1949*, National Electrical Manufacturers Association, New York, N. Y., Oct. 1949.

4. STANDARDS FOR INDUSTRIAL CONTROL. *NEMA Publication No. IC 1–1949*, National Electrical Manufacturers Association, New York N. Y., Nov. 1949.

5. IES LIGHTING HANDBOOK. Illuminating Engineering Society, New York, N. Y., 2nd ed., 1952.

6. ELECTRIC DISTRIBUTION AND CONTROL FOR LIGHTING SYSTEMS, W. H. Kahler, R. N. Bell. *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 92–99.

7. HIGHER LIGHTING VOLTAGES SAVE COPPER, F. C. Winkler, H. G. Barnett. *Electrical World*, New York, N. Y., Sept. 4, 1943, pp. 72–74.

8. 240/416-VOLT 3-PHASE 4-WIRE POWER AND LIGHTING SUPPLY FOR MODERN INDUSTRIAL PLANTS, William Shuler. *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 71–81.

9. ECONOMICS OF VARIOUS SECONDARY VOLTAGES FOR COMMERCIAL AREAS, T. C. Duncan, J. P. Neubauer, J. M. Comly, R. F. Lawrence, Miles Maxwell. *AIEE Special Publication S-66*, Sept. 1954, pp. 7–20.

10. DISTRIBUTION EQUIPMENT USED ON 265/460-VOLT NETWORKS AND ITS OPERATING FEATURES, L. Brieger, C. P. Xenis, A. J. Visson, J. DeLellis. *AIEE Special Publication S-66*, Sept. 1954, pp. 20–26.

11. THE RELATIVE FEASIBILITY OF 460-VOLT AND 208-VOLT SERVICE IN COMMERCIAL BUILDINGS, H. G. Barnett, R. A. Zimmerman, H. E. Lokay. *AIEE Special Publication S-66*, Sept. 1954, pp. 37–44.

---

480 Wye/277-Volt Power System in Telephone Building at Menands, N. Y.

D. S. BRERETON
ASSOCIATE MEMBER AIEE

H. J. DONNELLY
ASSOCIATE MEMBER AIEE

Synopsis: The Edison Electric Institute—National Electrical Manufacturers Association standard nominal system voltage designation of the higher voltage system referred to in this paper is 480 wye/277 volts. The maximum system voltage is 480 wye/277 and the average utilization voltage will be more nearly 460 wye/265 volts. To be consistent with the lower class designation, such as 208 wye/120 volts which is an average utilization voltage, some utilities have selected 460 wye/265 volts as a nominal system voltage designation. Whether the higher voltage system is called 480 wye/277 or 460 wye/265 volts, it is the same system and has the same voltage spread.

THERE is an increasing tendency for the modern commercial building to employ voltages above the conventional 208 wye/120 volts. This has been increasingly true for new buildings with a load of a few thousand kva or above. For these buildings, the distribution voltage has been 480 wye/277 volts with a substantial portion of the utilization at this voltage. These loads include lighting, air-conditioning motors, elevator and other service machinery, and motor-pump sets.

A small building, with a load of less than 1,000 kva, has proved to be more economical at 480 wye/277 than at 208 wye/120 volts in a number of cases. One such case is the Twin Falls High School at Twin Falls, Idaho.[1]

The New York Telephone Company was very interested in reviewing the cost comparison between the 480 wye/277 and 208 wye/120 voltage levels as well as a detailed analysis of the components that make up the complete power system for their new building at Menands, N. Y. For this office building, power is supplied at 4,160 volts by the utility, Fig. 1 bottom, and is transformed to 480 wye/277 volts through two 300-kva load center unit substations. A careful study would reveal the relative importance of the components that make up the system and would permit a close watch on items comprising the substantial parts of it. Such a study was made, and it was clearly evident that the 480 wye/277 voltage was the most economical voltage level. The New York Telephone Company and the General Electric Company, later reviewing the first study, concluded that the results of a still more detailed comparison would be of great value to the building industry. At this time the two companies agreed to share equally in the preparation of another set of complete drawings for the building, utilizing 208 wye/120 volts. It is the authors' opinion that such a complete comparison for an office building has thus far not been attempted.

The purpose of this paper is to make available the results of the economic studies at the two voltage levels and to show the effect of conventional 120-volt switching in comparison with 24-volt remote control branch-circuit switching. Two conclusions resulted: a revaluation of the branch-circuit wiring to the effect that it is a relatively minor part of the system for the building; the flexibility of a master control switching system takes on more importance than first realized.

Description of Building

The rectangularly shaped, brick veneered office building is 102.5 feet by 248.5 feet. With each floor having 25,470 square feet gross, the total area for the building becomes 76,410 square feet gross. The three floors are similar in design, the individual offices being provided with movable metal partitions. Fluorescent lighting is supplied by 8-foot

---

Paper S-66, recommended by the AIEE Transmission and Distribution and Industrial Power Systems Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 21, 1954; made available for printing July 14, 1954.

D. S. BRERETON is with the General Electric Company, Schenectady, N. Y., and H. J. DONNELLY is with the New York Telephone Company, Albany, N. Y.

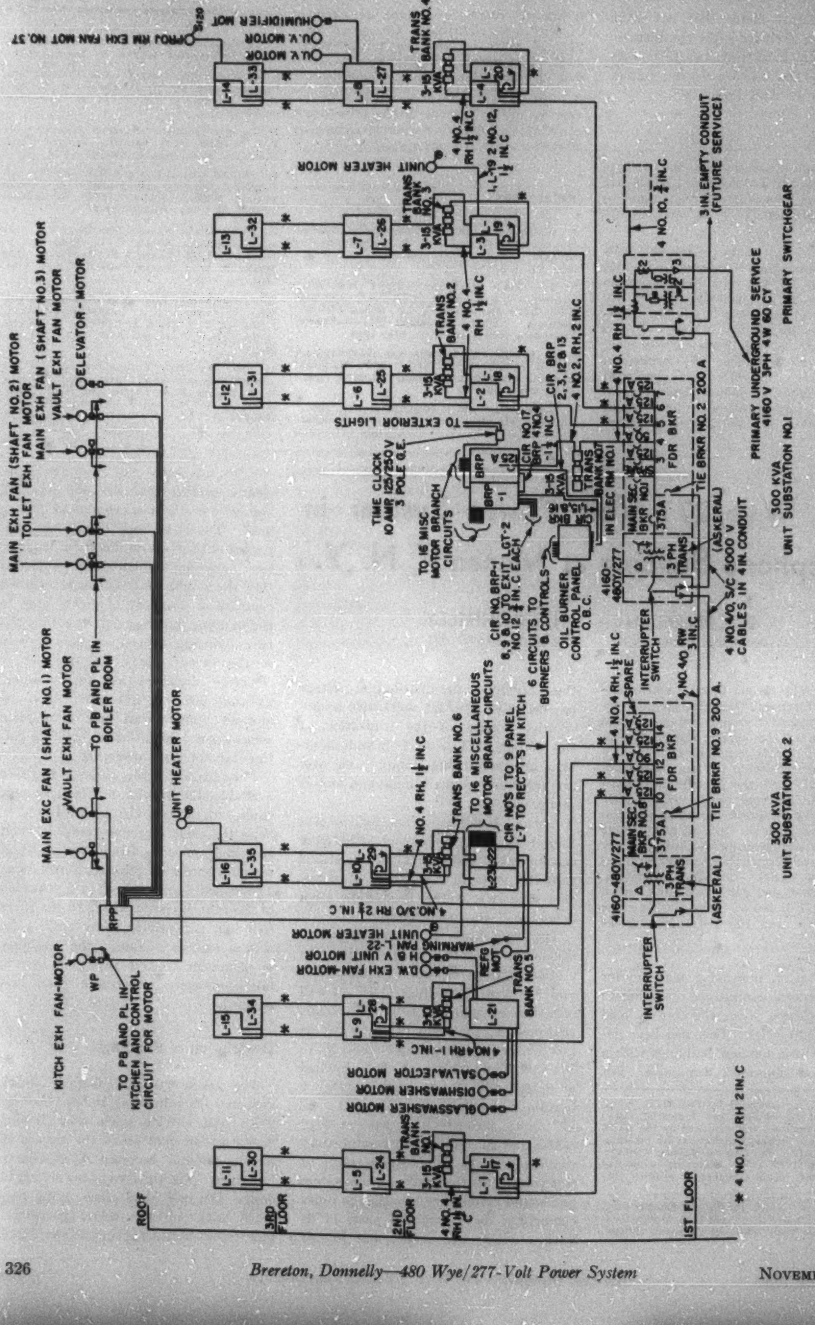

ROOF
3RD FLOOR
2ND FLOOR
1ST FLOOR
L-11
L-30
L-5
L-24
L-1
L-17
TRANS BANK NO.1
3-15 KVA
4 NO.4 RH 1½ IN.C
GLASSWASHER MOTOR
DISHWASHER MOTOR
SALVAJECTOR MOTOR
4 NO.4 RH 1-IN.C
L-15
L-34
L-9
L-28
3-10 KVA
L-21
TRANS BANK NO.5
REFG MOT
D.W. EXH FAN-MOTOR
H & V UNIT MOTOR
WARMING PAN L-22
UNIT HEATER MOTOR
4 NO.3/0 RH 2½ IN.C
KITCH EXH FAN-MOTOR
WP
TO PB AND PL IN KITCHEN AND CONTROL CIRCUIT FOR MOTOR
RPP
L-16
L-35
L-10
L-29
L-23-22
CIR NO'S 1 TO 9 PANEL L-7 TO RECPT'S IN KITCH
TRANS BANK NO.6
TO 16 MISCELLANEOUS MOTOR BRANCH CIRCUITS
4 NO.4 RH, 1½ IN.C
UNIT HEATER MOTOR
MAIN EXC FAN (SHAFT NO.1) MOTOR
VAULT EXH FAN MOTOR
TO PB AND PL IN BOILER ROOM
MAIN EXH FAN
TOILET EXH FAN MOTOR
MAIN EXH FAN (SHAFT NO.2) MOTOR
MAIN EXH FAN (SHAFT NO.3) MOTOR
VAULT EXH FAN MOTOR
ELEVATOR-MOTOR
L-12
L-31
L-6
L-25
L-2
L-18
TRANS BANK NO.2
4 NO.4 RH 1½ IN.C
TO EXTERIOR LIGHTS
TIME CLOCK 10 AMP 125/250 V 3 POLE G.E.
BRP
BRP-1
TO 16 MISC MOTOR BRANCH CIRCUITS
CIR NO. BRP-1 6,8 & 10 TO EXIT LGT-2 NO.12 ¾ IN.C EACH
6 CIRCUITS TO BURNERS & CONTROLS
OIL BURNER CONTROL PANEL O.B.C.
CIR BKR 15 & 16
CIR NO 17 BRP 4 NO.4 1¼ IN.C
CIR BRP 2, 3, 12 & 13 4 NO.2, RH, 2 IN.C
TRANS BANK NO.7 IN ELEC RM NO.1
L-13
L-32
L-7
L-26
L-3
L-19
TRANS BANK NO.3
1, L-19 2 NO.12, 1½ IN.C
UNIT HEATER MOTOR
4 NO.4 RH 1½ IN.C
L-14
L-33
L-8
L-27
L-4
L-20
TRANS BANK NO.4
PROJ RM EXH FAN MOT NO.37
U.V. MOTOR
U.V. MOTOR
HUMIDIFIER MOT
NO.4 RH 1½ IN.C
4160-480Y/277
MAIN SEC BKR NO.1
375A
3 PH TRANS
FDR BKR
TIE BKR NO.2 200 A
(ASKERAL)
INTERRUPTER SWITCH
4 NO.4/0 S/C 5000 V CABLES IN 4 IN. CONDUIT
4 NO.4/0 RW 3 IN.C
300 KVA UNIT SUBSTATION NO.1
PRIMARY UNDERGROUND SERVICE 4160 V 3PH 4W 60 CY
PRIMARY SWITCHGEAR
4 NO.10, ¾ IN.C
3 IN. EMPTY CONDUIT (FUTURE SERVICE)
4 NO.4 RH, 1½ IN.C
SPARE
4160-480Y/277
MAIN SEC BKR NO.8
375A
3 PH TRANS
FDR BKR
TIE BKR NO.9 200 A.
(ASKERAL)
INTERRUPTER SWITCH
300 KVA UNIT SUBSTATION NO.2
* 4 NO. 1/0 RH 2 IN.C

Fig. 1. One-line and riser diagram for the 480 wye/277-volt system

Table I. Loads for 480-Volt System

	Kva	Per Cent
120-Volt		
Kitchen equipment	18	
Boiler room equipment	30	
Lighting incandescent	40	
Office machines (estimated)	182	
Total Kva	270	54.6
480-Volt		
Elevator	10	
Ventilating motors	26	
Total kva	36	7.2
480 Wye/277-Volt Lighting		
Total kva	190	38.2
Building kva	496	100.0

slimline tubes. Incandescent lighting, operating at 120 volts, is used in the cafeteria and other isolated areas. Fig. 2 shows the exterior of the building.

Basis for System Studies

In choosing the 480 wye/277-volt power system for the new telephone building, a very careful review was made of the economic and engineering features of supplying power by using: 1. a 480 wye/277-volt system with 24-volt remote control switching, 2. a 208 wye/120-volt system with 24-volt remote control switching; and 3. a 208 wye/120-volt system with regular 120-volt switchings.

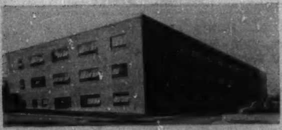


Fig. 2. New York Telephone Company office building at Menands, N. Y.

Fig. 3. Load center unit substation rated 300 kva, 4,160-480 wye/277 volts. The 4.16-kv metal-clad switchgear is to the right

One of the first steps in comparing the choice of using a secondary voltage level of 480 wye/277 or 208 wye/120 volts is to itemize the loads in the building. Because the loads would all be at 120 volts within a 208 wye/120-volt system, with the exception of the elevator and ventilating motors which would be 208 volts, 3 phase, the loads are itemized in Table I for the 480-volt system only.

The lighting load represents about 3.0 volt-amperes per square foot, including both the fluorescent and incandescent lighting, against the 120-volt office machines at 2.4 volt-amperes per square foot, and all the remaining loads in the building at 1.1 volt-amperes per square foot.

Based on past generalizations, it may be thought that a 480 wye/277-volt system would not be more economical than a 208 wye/120-volt system for a small building such as this in which more than half the load must operate at 120 volts. The general application of a 480-volt system for small buildings is discussed in more detail in the conclusions of the paper. The results of the comparisons that follow show that the 480 wye/277-volt distribution can be more economical for a small building with only a few integral-horsepower machines and having a 120-volt load as high as 54.6 per cent of the total.

Note that this telephone building does not include air-conditioning equipment of any kind. If it had been included as a centralized air-conditioning equipment, it would have increased the load and could have used motors rated 440 volts. This would most certainly make the 480 wye/277-volt system even more economical because the percentage of 480-volt load would be higher.

The present illumination level in the building averages 40 foot-candles. Provision to increase the lighting level in any area has been made by equipping each fixture with a knockout at the end for lamp socket mounting, and space in the wiring channel for the addition of a second series lamp ballast, thus providing the ability to add a third lamp. The lighting level in the building is based on a maintenance factor of 70 per cent.

$$\text{Foot-candles} = \frac{\text{(lumens per lamp)(no. of lamps)(coefficient of utilization)(maintenance factor)}}{\text{Area in square feet}}$$

For this building, to obtain "seeing comfort," a 45-degree by 45-degree shielding was selected. The louvres to produce this shielding are steel "egg crates" with white baked-on enamel.

Fig. 4. Panelboard layout conduit for branch-circuit wiring and connections to the underfloor duct system

Louvre brightness was controlled by employing matte aluminum reflectors.

Installed Costs

Although the cost of materials, such as lighting fixtures, cables, or substations, is lowest for the 480 wye/277-volt system, an evaluation based on equipment cost only would be inconclusive for a study such as this. A preferred and a more representative comparative base is installed cost. Previous studies[2] were also made on

Fig. 5. Typical panelboard layout during construction. The electric services are in the box at the left and the telephone and signal circuits are handled in the two boxes to the right

the basis of installed cost. These studies omitted items related to the electric system such as the underfloor duct system, lights and fixtures, and the actual first cost of machines, but did include the effects of such items. This paper gives the additional information for the sake of completeness.

The system cost comparisons are based on sound system engineering. It should be noted, however, that local codes, type of building construction, and other factors may alter the engineering choice of equipment and the total installed cost for a similar building in other localities. One of the National Electrical Code requirements enforced for this building was the required use of asbestos wire to make direct connection to the fluorescent fixtures. This and other points dealing with codes are covered in detail later.

Incoming 4,160-Volt Service

A single 4,160-volt service was made available from the local utility. Their service terminated at the utility pole and the telephone company supplied the cable service to their metal-clad switchgear equipment, shown in Fig. 3, which consisted of an entrance compartment for the power company's metering, a second compartment containing the control power transformer, and a third compartment containing a single magne-blast power circuit breaker, rated 4.16 kv, 150 megavolt-amperes interrupting, 1,200 amperes, and using a-c close and d-c trip from a 24-volt battery.

Four single-conductor no. 4/0 5-kv Super Coronol* cables are used to connect between the metal-clad switchgear and the unit substations; see bottom of Fig. 1. These services are terminated at the load center primary Pyranol* filled switches. Provision has been made for connecting additional substations to this 4,160-volt cable service. The fourth wire, neutral ground wire, is terminated on the ground pad of the load center substation transformer. The neutral is presently disconnected at the utility pole. It is hoped that the connection of this neutral wire to the utility system will soon be made. This will decrease the zero phase sequence impedance to assure a proper and low impedance ground return of ground-fault current through this return neutral wire instead of forcing it through a building structure or piping, should a fault occur in one of the transformers. A section of the metal-clad switchgear, along with one of the 300-kv load center unit substations, is shown in Fig. 3.

Fig. 6. Electric services in the panelboard on the left. This panelboard has 277-volt fused circuits (left), transformer secondary breaker for 120-volt transformers (top), and 120-volt circuits (right)

The two load center unit substations are located on the first floor and have a normally open secondary tie, bottom of Fig. 1, to form a "secondary selective" system. This type of system is very adaptable to commercial buildings supplied at primary voltage and is being in-

---

* Registered trademark of General Electric Company.

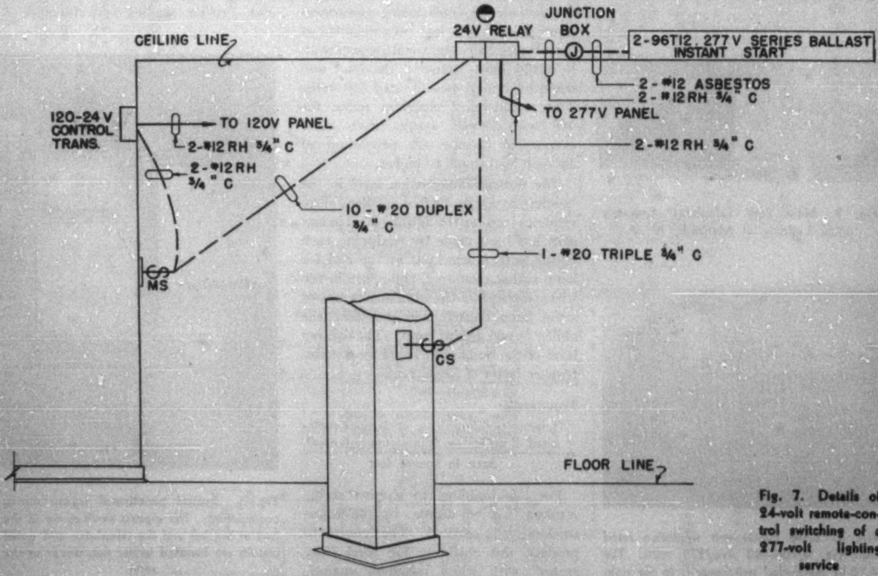


Fig. 7. Details of 24-volt remote-control switching of a 277-volt lighting service

creasingly specified. This arrangement permits one of the substation transformers to be removed from service for routine maintenance. By closing the tie circuit breaker, the other transformer can supply the essential loads on both substation busses to keep power available everywhere in the building. This system provides a high order of safety because the switching, for transferring load, is handled on the secondaries by an adequate circuit breaker.

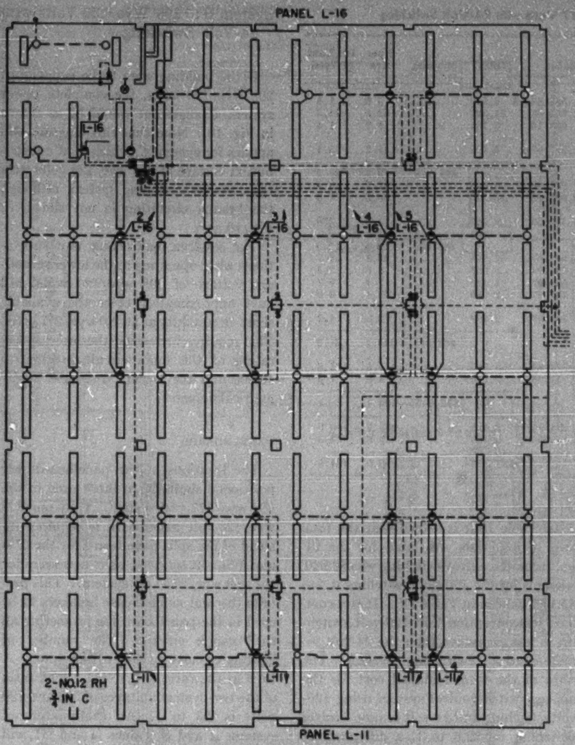


Fig. 8. 480 wye/277-volt system with 24-volt switching

Fig. 9. Ceiling of one section showing the 277-volt circuiting to the conduit box for a fluorescent fixture and the split conduit box for the 24-volt control

System A: 480 Wye/277 Volts with 24-Volt Switching

The complete 1-line drawing of this power system arrangement is shown in Fig. 1, giving such additional details as the complete rating of the incoming and feeder cable services, the method of feeding the panelboards, and the details of the motors and their services. The 300-kva load center unit substation, rated 480 wye/277 volts secondary, uses a main circuit breaker rated 25,000 and feeder breakers rated 15,000 asymmetrical amperes interrupting.

PANELBOARDS

Convenience necessitated the use of two panelboards to serve one-third of each floor. The telephone company required each panelboard location on a floor to be served by one feeder. Therefore, a separate feeder serves one panelboard on each of the three floors, Fig. 4, or three panelboards per feeder. This required each of the two substations to have three feeder breakers for lighting and office machines, and one feeder for special services.

The various panelboards making up the panelboard layout are shown in Figs. 5 and 6. The composite Trumbull panelboard on the left of the layout contains the switching for the electric services. This panelboard contains three parts: 1. the right-hand section houses the small 120-volt molded-case breakers for protecting the 120-volt circuits; 2. the bottom section houses the molded-case breaker on the secondary side of the three single-phase dry-type *M*-transformers—generally three 15-kva units—which feed the 120-volt molded-case breakers; and 3. the left-hand section houses a Swingwa* panelboard which uses silver-sand current-limiting fuses. This fuse-switch combination has been successfully tested to 50,000 asymmetrical amperes short-circuit current.

---

* Registered trademark of General Electric Company.

Table II. System A: 480 Wye/277 Volts with 24-Volt Switching

	Material	Labor	Total	Subtotal	$ per Kva	Per Cent of Total
Primary feeders	$ 2,393.72	$ 581.88	$ 2,976		5.0	1.4
Two 300-kva substations	28,572.24	2,754.50	31,327		52.2	14.3
Secondary feeders	2,083.52	1,051.15	3,134		5.2	1.4
Fifteen 15-kva and three 10-kva 480–120-volt transformers	5,474.02	1,267.07	6,741		11.2	3.1
				$44,178	73.6	20.2
Panelboards	8,871.00	1,731.89	10,603		17.7	4.9
Lighting branch circuit	5,256.11	10,063.73	15,320		25.5	7.0
Receptacle, etc., branch circuits, 120 volts	3,445.60	4,510.50	7,956		13.3	3.6
				$33,879	56.5	15.5
Underfloor duct system	46,145.94	10,921.61	57,068		95.1	26.1
Motor wiring and equipment	3,398.05	1,694.94	5,093		8.5	2.3
Telephone and signal system	727.00	688.63	1,416		2.4	0.7
Telephone and signal cabinets	1,801.91	237.57	2,039		3.4	0.9
Miscellaneous, hangers, pullboxes, etc.	2,250.68	346.97	2,598		4.3	1.2
				$68,214	113.7	31.2
Fluorescent and incandescent fixtures and lamps	61,031.71	8,273.84	69,305		115.5	31.8
Column switch control	1,228.48	1,711.66	2,940		4.9	1.3
				$72,245	120.4	33.1
Total	$172,679.98	$45,835.94	$218,516		364.2	100.0
Master switch control, 24 volts	3,570.68	5,889.09	9,460		15.8	4.3
			$227,976		380.0	104.3

277-VOLT BRANCH CIRCUITS

Fig. 7 shows a typical 277-volt branch circuit illustrating how connections were made from the panelboard to the conduit boxes serving the fluorescent fixtures. Fig. 8 illustrates how the branch circuits for one-third of a floor, terminate at special conduit boxes which contain the 24-volt control relays, denoted by a half-dark and half-light circle. The power circuits are then connected to other conduit boxes, a light circle, which directly connect to the fluorescent fixtures. This is also illustrated in Fig. 9, showing the ceiling before completion and both of these conduit boxes.

24-VOLT REMOTE-CONTROL SWITCHING

The 24-volt remote-control switching circuits are designated by the light dashed line in Fig. 8. Fig. 7 shows these switching circuits, including the 24-volt circuit to the wall switch *CS* and the 24-volt circuit to the master control switching center *MS*. It is easily seen that this system includes a transformer to supply control power at 24 volts for use on the control switches, remote-control relays, and small 24-volt wall switches for each circuit. As well as permitting only 24 volts to be within the reach of the building occupants, this system offers outstanding advantages for office buildings, since it is extremely flexible, convenient to install, safe to operate, and makes a substantial saving in the over-all system cost possible.

In Table II it is shown that the total cost of materials and labor for the 24-volt remote-control switching was $2,940, whereas for the 120-volt switching it was $5,112, as seen in Table IV. It is a common misconception that 120-volt switching is less expensive than the 24-volt remote-control switching system. The ratio of the total installed cost for the master switch control system using 120-volt switching and 24-volt remote-control switching was 2.6 to 1, a difference of over $13,000.

System B: 208 Wye/120 Volts with 24-Volt Switching

If the building were to be supplied at 208 wye/120 volts, the complete power system arrangement would be as shown in Fig. 10. Note that this diagram also applies for system *C*, which is at a distribution voltage of 208 wye/120 volts with the 120-volt switching system, although the branch circuiting is not shown in Fig. 10.

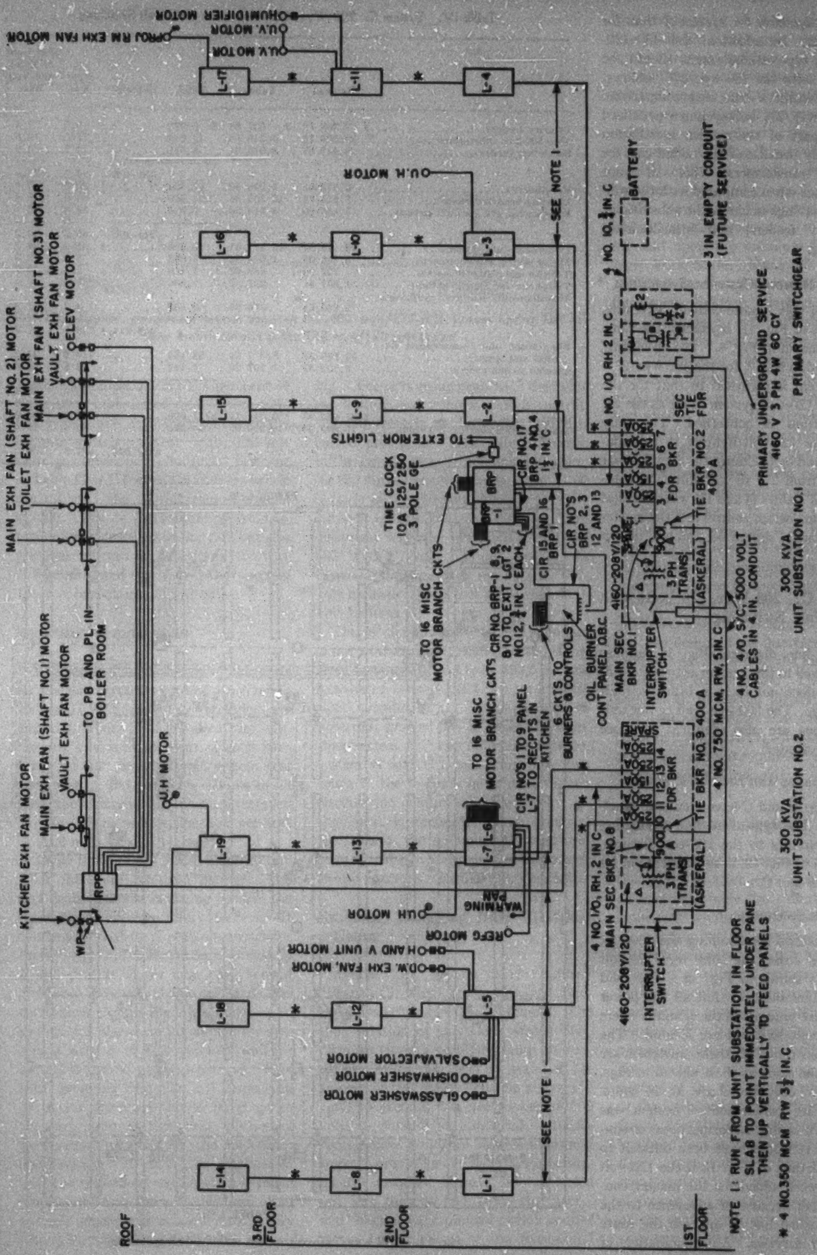
The 300-kva load center unit substations, when operating at the lower secondary voltage of 208 wye/120 volts, will have approximately twice the available short-circuit duty as at 480 wye/277 volts. This requires increasing the interrupting rating of the main circuit breaker to 50,000 and the feeders to 25,000 asymmetrical amperes.

PANELBOARDS

The locations of the panelboards are positioned similarly to those used in the 480 wye/277-volt system. Only one 208 wye/120-volt panelboard is required instead of the split panelboard for the 277- and 120-volt services, as is necessary for the 480 wye/277-volt system. This permits thermal molded-case breakers to be used in the panelboard for protecting all the branch circuits. The number of lighting circuits increases from system *A* to *B* in the ratio of 2.1 to 1. The ratio of the two system voltages of 277 and 120 volts is 2.3 to 1. An examination of systems *A* and *B*, Tables II and III, will immediately show that the panelboards

Table III. System B: 208 Wye/120 Volts with 24-Volt Switching

	Material	Labor	Total	Subtotal	$ per Kva	Per Cent of Total, 480-Volt Base
Primary feeders	$ 2,393.72	$ 581.88	$ 2,976		5.0	1.4
Two 300-kva substations	39,698.15	3,213.59	42,912		71.5	19.6
Secondary feeders	6,441.67	2,076.21	8,518		14.2	3.9
				$54,406	90.7	24.9
Panelboards	9,925.65	2,398.90	12,324		20.5	5.6
Lighting branch circuits	6,390.00	12,210.00	18,600		31.0	8.5
Receptacles, etc., branch circuits	3,446.60	4,510.50	7,956		13.3	3.7
				$38,880	64.8	17.8
Underfloor duct system	46,145.94	10,921.61	57,068		95.1	26.1
Motor wiring and equipment	3,398.05	1,694.94	5,093		8.5	2.3
Telephone and signal system	727.00	688.63	1,416		2.4	0.7
Telephone and signal cabinets	1,801.91	237.57	2,039		3.4	0.9
Miscellaneous, hangers, pullboxes, etc.	2,250.68	346.97	2,598		4.3	1.2
				$68,214	113.7	31.2
Fluorescent and incandescent fixtures and lamps	59,890.59	8,273.84	68,164		113.6	31.2
Column switch control	1,228.48	1,711.66	2,940		4.9	1.3
				$71,104	118.5	32.5
Total	$183,738.44	$48,866.30	$232,604		387.7	106.4
Master switch control, 24 volts	3,570.68	5,889.09	9,460		15.8	4.3
			$242,064		403.5	110.7


ROOF
3RD FLOOR
2ND FLOOR
1ST FLOOR
L-1
L-2
L-3
L-4
L-5
L-6
L-7
L-8
L-9
L-10
L-11
L-12
L-13
L-14
L-15
L-16
L-17
L-18
L-19
RPP
WP
BRP
BRP-1
SEE NOTE 1
SEE NOTE 1
PROJ RM EXH FAN MOTOR
HUMIDIFIER MOTOR
U.V. MOTOR
U.V. MOTOR
U.H. MOTOR
U.H. MOTOR
U.H MOTOR
MAIN EXH FAN (SHAFT NO.2) MOTOR
TOILET EXH FAN MOTOR
MAIN EXH FAN (SHAFT NO.3) MOTOR
VAULT EXH FAN MOTOR
ELEV MOTOR
KITCHEN EXH FAN MOTOR
MAIN EXH FAN (SHAFT NO.1) MOTOR
VAULT EXH FAN MOTOR
TO PB AND PL IN BOILER ROOM
TO EXTERIOR LIGHTS
TIME CLOCK 10A 125/250 3 POLE GE
TO 16 MISC MOTOR BRANCH CKTS
TO 16 MISC MOTOR BRANCH CKTS
CIR NO. BRP-1 8, 9, & 10 TO EXIT LGT 2 NO.12, 3/4 IN C EACH
CIR NO.17 BRP 4 NO.4 1½ IN. C
CIR 15 AND 16 BRP 1
CIR NO'S BRP 2, 3 12 AND 13
6 CKTS TO BURNERS & CONTROLS
OIL BURNER CONT PANEL O.B.C
CIR NO'S 1 TO 9 PANEL L-7 TO RECPTS IN KITCHEN
WARMING PAN
REFG MOTOR
H AND V UNIT MOTOR
D.W. EXH FAN. MOTOR
SALVAJECTOR MOTOR
DISHWASHER MOTOR
GLASSWASHER MOTOR
4 NO. 1/0 RH 2 IN. C
NO. 10, 3/4 IN. C
BATTERY
3 IN. EMPTY CONDUIT (FUTURE SERVICE)
PRIMARY UNDERGROUND SERVICE 4160 V 3 PH 4W 60 CY
PRIMARY SWITCHGEAR
SEC TIE FDR
TIE BKR NO. 2 400 A
FDR BKR
4160-208Y/120
MAIN SEC BKR NO.1
INTERRUPTER SWITCH
3 PH TRANS (ASKERAL)
4 NO. 4/0, S/C, 5000 VOLT CABLES IN 4 IN. CONDUIT
4 NO. 750 MCM, RW, 5 IN. C
TIE BKR NO. 9 400 A
4 NO. 1/0, RH, 2 IN C MAIN SEC BKR NO.8
4160-208Y/120
INTERRUPTER SWITCH
3 PH TRANS (ASKERAL)
SPARE
225A
150A
100A
300 KVA UNIT SUBSTATION NO.1
300 KVA UNIT SUBSTATION NO.2

NOTE 1: RUN FROM UNIT SUBSTATION IN FLOOR SLAB TO POINT IMMEDIATELY UNDER PANE THEN UP VERTICALLY TO FEED PANELS

* 4 NO.350 MCM RW 3½ IN.C

Fig. 10. One-line and riser diagram for the 208 wye/120-volt system

are less expensive for system *A* than for system *B*. In addition, the 480–120-volt dry-type transformers should be included with the 480 wye/277-volt system to obtain a true cost comparison. The 480 wye/277-volt system is penalized in this part of the system installation principally because of the added cost for dry-type transformers. This is of minor importance when compared with the substantial savings made in the substations, secondary feeders, and branch-circuit switching.

24-Volt Remote-Control Switching

The 24-volt remote-control switching, as shown in Fig. 11, is exactly the same as that shown in Fig. 8. The circuiting, illustrated in Fig. 7, for the 480 wye/277-volt system is identical to that for the 208 wye/120-volt system as far as the 24-volt portion is concerned. The only change for 120-volt lighting in this diagram would be to change the actual power services from 277 to 120 volts. A comparison of Tables II and III will show that the column switch control and the master switch control should be, and are, identical.

System C: 208 Wye/120 Volts with 120-Volt Switching

The 1-line diagram, Fig. 10, applies equally well to this system as to system *B*, since it does not include branch-circuit switching. The 300-kva load center unit substations are also identical to those used in system *B*.

Panelboards and Branch Circuits

Systems *B* and *C* have identical panelboards. The typical method of running branch circuits to the 120-volt fixtures, as shown in Fig. 12, is again identical for systems *B* and *C*.

120-Volt Switching

The method for handling the 120-volt switching from the 208 wye/120-volt system is shown in Fig. 13. The solid dashed lines indicate that all the power circuits are brought to the columns where the 120-volt switches are located. The actual locations of these switches are exactly the same as those shown in Figs. 8 and 11, only those are at 24 volts. The master switch control system was not shown in Fig. 13 for purposes of simplicity. It would have been difficult to see the details of Fig. 13 if the 120-volt circuits were added for the master control system because of adherence to the heavy dashed line to indicate the route of power services. An examination of

Table IV. System C: 208 Wye/120 Volts with 120-Volt Switching

	Material	Labor	Total	Subtotal	$ per Kva	Per Cent of Total, 480-Volt Base
Primary feeders	$ 2,393.72	$ 581.88	$ 2,976		5.0	1.4
Two 300-kva substations	39,698.15	3,213.59	42,912		71.5	19.6
Secondary feeders	6,441.67	2,076.21	8,518		14.2	3.9
				$54,406	90.7	24.9
Panelboards	9,925.65	2,398.90	12,324		20.5	5.6
Lighting branch circuits	8,215.74	13,304.34	21,520		35.9	9.9
Receptacles, etc., branch circuits	3,446.60	4,510.50	7,956		13.3	3.7
				$41,800	69.7	19.2
Underfloor duct system	46,145.94	10,921.61	57,068		95.1	26.1
Motor wiring and equipment	3,398.05	1,694.94	5,093		8.5	2.3
Telephone and signal system	727.00	688.63	1,416		2.4	0.7
Telephone and signal cabinets	1,801.91	237.57	2,039		3.4	0.9
Miscellaneous, hangers, pullboxes, etc.	2,250.68	346.97	2,598		4.3	1.2
				$68,214	113.7	31.2
Fluorescent and incandescent fixtures and lamps	59,890.59	8,273.84	68,164		113.6	31.2
Column switch control	3,120.23	1,992.26	5,112		8.5	2.4
				$73,276	122.1	33.6
Total	$187,455.93	$50,241.24	$237,696		396.2	108.9
Master switch control, 120 volts	13,849.00	10,980.00	24,829		41.4	11.3
				$262,525	437.6	120.2

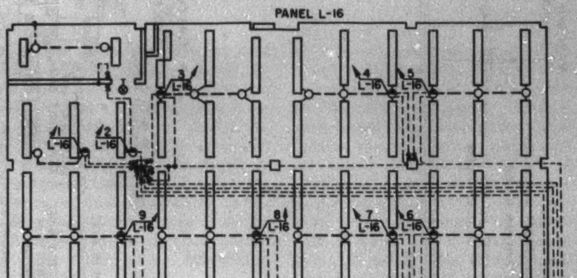


Fig. 11. 208 wye/120-volt system with 24-volt switching

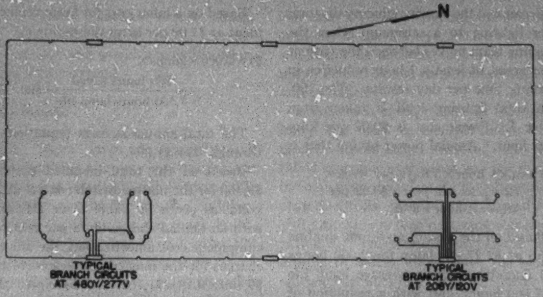


Fig. 12. Typical panelboard layout showing the 480 wye/277-volt system having half as many branch circuits as the 208 wye/120-volt system

Table IV indicates that 2.4 per cent of the system cost, based on 480 wye/277-volt system cost as 100 per cent, is in the column switch control with 120-volt switching whereas it is only 1.3 per cent in Table II or III when the column switch control used the 24-volt remote-control system. The 24-volt remote-control system not only costs less in this case but also provides substantial benefits in the way of flexibility and safety over the common 120-volt branch-circuit switching.

Economic Comparisons

The detailed breakdowns of the costs for the three systems studied are given in Tables II, III, and IV. Included are material, labor, total installed cost, dollars per kva, and per cent of each item using the 480 wye/277-volt system cost as a base of 100 per cent. Items in all three tables having the identical installed cost, in dollars, will be the same per cent of total cost.

These three tables group material and labor into four subtotal listings. The first deals with the incoming service, the substations, and secondary feeders. It can be seen that there is an advantage for the 480 wye/277-volt system over system *B* and *C* of $17.1 per kva. The second grouping, including the panelboards and the branch circuits for both lighting and receptacles, shows an advantage for system *A* over *B* of $8.3 per kva, and for system *A* over *C* of $13.2 per kva. The third grouping lists identical items, and particular attention should be given to the very large, 26.1-per-cent, portion of the system required for the underfloor duct system. It should be emphasized that the underfloor duct system is much less expensive in the long run than other systems where it is necessary to modify branch-circuit wiring in the floor or ceiling. The next group deals with the lighting fixtures and column switch control. The difference in the latter has been discussed previously. The only difference in the two figures, at 277 and 120 volts, for the fluorescent lighting is about $1.60 in favor of the 120-volt ballast. It is the authors' opinion that this difference will disappear with the increased use of the 480 wye/277-volt lighting system, as there is no engineering reason why 277-volt ballasts should cost more than 120-volt ballasts.

The most important item to compare is the total cost for all material and labor. Here system *A* proves to be $23.5 per kva less than system *B* and $32.0 per kva less than system *C*. It is important to note that system *C* is $8.5 per kva more than system *B*. The difference between system *B* and *C* comes from only two sources; the increase in lighting branch circuits and column switch control. Although it is certainly not a practical comparison, the difference between systems *A* and *C*, when including the master switch control system, is $57.6 per kva.

Problems Resulting from Electrical Codes

Article 210, section 2113, entitled "Branch Circuits," of the National Electrical Code[3] deals with "Voltage." The 1953 revision of this Code states that, "Branch circuits supplying lamp holders, fixtures, or receptacles of the standard 15-ampere or less rating" can be above 150 volts to ground, but not to exceed 300 volts to ground "in industrial establishments, office buildings, large schools and stores," providing the ballasts for permanently installed fixtures be "mounted not less than 8 feet from the floor," and "do not have manual switch control as an integral part of the fixture(s)." This requirement of the National Electrical Code has been completely adhered to in the telephone building.

Additional safety has been provided in this building, beyond the requirements of most codes, in that all the branch-circuit switching throughout the building, whether the line-to-neutral voltage is 277 or 120 volts, is handled by 24-volt switches.

The New York Telephone Company has had extensive experience in constructing and operating buildings throughout New York State. Many of these new buildings have used recessed fixtures for fluorescent lighting. Section 94105 of the Code[3] deals with "Flush and Recessed Fixtures" and requires, under item "f," "Conductor Insulation," that the "Recessed fixtures shall be wired with AF or AI fixture wire." The local inspector had always waived this requirement for recessed fluorescent fixtures provided adequate means for cooling the fixture are incorporated in the ceiling construction. An identical construction was employed in this building as in many others, but because this was the first building to apply the 480 wye/277-volt lighting system in an office-type building in the Albany area it was believed that conservatism demanded that this requirement be fulfilled exactly. It is appropriate to add that the design of the fluorescent fixtures prohibits easy access to the fixture by its dead-front construction. The vast experience in industrial plants[4] with 277-volt lighting indicates that the added expense of using asbestos wire, at either 277 or 120 volts, is not required from a safety or engineering standpoint.

Remote-Control System

The building maintenance foreman at this building, after less than two weeks' observation, made the statement, "This is the only building I ever heard of where the lighting control system was designed with the maintenance people in mind." A source of constant irritation for maintenance people is the recurring observation by top management that buildings use too much light at night. The management point is always well taken, but the means for effective and economical control of the problem is seldom placed in the maintenance department's hands. In this installation, master switch control has been furnished at three locations per floor. Each half-floor has complete remote control from the north and south staircases and each full floor is controlled at the center staircase. Therefore, it is easy to turn all the lights off and on from these locations without walking over the entire floor. The following text gives a cost estimate, based on limited experience, which substantiates the preceding statement.

Group or team cleaning is employed in the building, i.e., a team finishes work in one area, for instance a half-floor, then moves to the next, in place of assigning a section to a particular cleaner. The time saving from centralized lighting control amounts to $1^1/_2$ man-hours per day. According to the maintenance department, they save $7^1/_2$ minutes on each of the six half-floors in turning the lights on and off.

Based on \$2.00 per man-hour for cleaning personnel and a saving in cleaning time of $1^1/_2$ man-hours per day, the annual labor saving is

$$1^1/_2 \text{ man-hours per day} \times \$2 \text{ per man-hour} \times 5 \text{ days per week} \times 52 \text{ weeks per year} = \$780 \qquad (1)$$

Because of the ease in reducing the total floor lighting to a minimum when the cleaning team moves between the various floor areas, at least a 1-hour reduction in burning time per day results. The 190-kva total lighting load is taken from Table I. Power cost is \$0.02 per kilowatt-hour. Annual power saving thus is

$$190 \text{ kva} \times 1 \text{ hour} \times 5 \text{ days per week} \times 52 \text{ weeks per year} \times \$0.02 \text{ per kilowatt-hour} = \$988 \qquad (2)$$

Based on the cost of an 8-foot slimline lamp at \$2.27 per lamp for 2,000 lamps (all floors) having an average lamp life of 7,500 hours, and with a reduction of burning time of 1 hour per day, for a 5-day week, or 260 hours per year, the annual lamp saving cost is

$$\$2.27 \text{ per lamp} \times 2{,}000 \text{ lamps} \times \frac{260 \text{ hours saved}}{7{,}500 \text{ hours lamp life}} = \$157 \qquad (3)$$

Based on a labor cost for lamp replacement at \$1.00 per lamp the saving is

$$\$1 \times 2{,}000 \times \text{lamps} \times \frac{260 \text{ hours saved}}{7{,}500 \text{ hours lamp life}} = \$69 \qquad (4)$$

The total annual savings (equations 1 through 4) = \$1,994.

Based on the total installed cost of \$9,460 for the master switch control at 24 volts, as given in Table II or III, and with an annual charge of 18 per cent for equipment cost amortization, the annual charges for the master switch control are \$9,460×0.18 = \$1,703. Therefore, the net savings are the total annual savings (\$1,994) minus the annual charges (\$1,703) or \$291.

Cost of Relay Conduit Boxes

Included in Tables II and III, under "lighting branch circuit," is the cost for individual conduit boxes housing the remote-control relays. It will be noted from Fig. 1 that these relays are located in the ceilings along the branch-circuit runs and near the column switches. The deletion of the 252 individual relay boxes and the relocation of the 252 relays in 18 relay boxes, one at each panelboard location, was also studied. The cost for this was 252 boxes×\$6.32 per box = \$1,593; 18 boxes×\$30 per box = \$540. Therefore the apparent saving = \$1,053.

The use of the 18 relay boxes would involve extension of the 24-volt wiring and conduits from the 252 ceiling relay locations to the panelboard locations, Fig. 12, at a cost equal to or exceeding the \$1,053 previously mentioned. Applying a motor-operated master control switching system, such as the General Electric "Circuit Servant" plan, to the 18 panelboard locations would simplify and reduce the cost

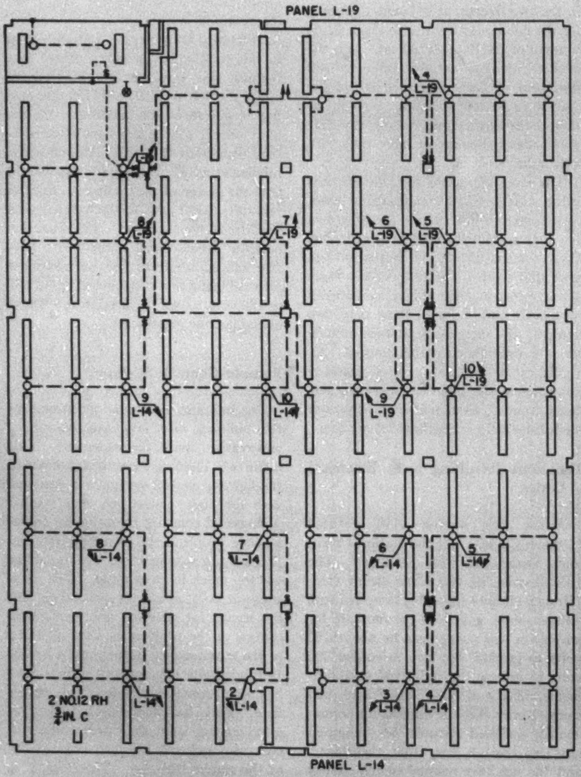


Fig. 13. 208 wye/120-volt system with 120-volt switching

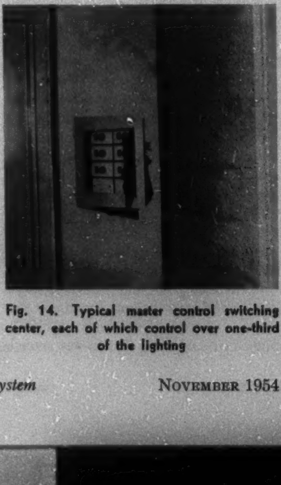


Fig. 14. Typical master control switching center, each of which control over one-third of the lighting

of the installed master switch wiring, but would add some $1,100 to the apparatus cost and would impair the selectivity of the system.

Conclusions

This paper gives a comprehensive evaluation and complete cost comparison between 208 wye/120- and 480 wye/277-volt power systems for an office building. It shows that even for this small building, with 54.6 per cent of its load at 120 volts, it is $23.5 per kva less at 480 wye/277 volts than at 208 wye/120 volts with comparable switching. It also shows a saving of $32.0 per kva between the 480 wye/277-volt system with 24-volt switching and 208 wye/120-volt system with 120-volt switching.

Now that 277-volt wall switches are available, it is possible to add a fourth comparison which would be a 480 wye/277-volt system with 277-volt switching. One thing is clearly evident about its application from the tables in this paper. Table II would be very similar and would only have variations in the lighting branch circuits and column switch control, as Table IV differs from Table III. The important point is that although the 24-volt switching would be more economical than 277-volt switching, the over-all economy of the 480 wye/277-volt system would not be affected adversely to any major extent by the employment of 277-volt switches.

Emphasis is again placed on the flexibility that must be incorporated in the modern commercial building. A detailed economic evaluation was not made in comparing the 24- and 120-volt master switch controls. One point in favor of 24-volt switching is that continual reorganization and reassociation of persons operating in offices, to improve the office efficiency, results in a great popularity of providing movable partitions used to form the office areas. The changes in switching control when at 24 volts can be made easily and in a comparatively inexpensive manner by simply running the control wire in the underfloor duct, within conduit or wire mold or even exposed. This is much easier to do than for the larger power carrying services. The cost of a lighting system is not only the initial cost but also includes the cost of making changes in the system when the future demands.

It is important to mention that the relative difference in dollars per kva between the 480 wye/277- and 208 wye/120-volt systems would be larger had this same size building been air-conditioned. Experience indicates that a building, in this locality, would approximately double its load if it were entirely air-conditioned. This air-conditioning equipment would use integral-horsepower motor rated 440 volts. The increase in load at 480 volts would produce a greater difference in dollars per kva between the 480 wye/277- and 208 wye/120-volt systems. Even if small unit air conditioners were used, the increase in motor load could more easily and economically be supplied from a 480-volt system. In fact, it can be said that as the load grows in the modern commercial building using fluorescent lighting more economical operation and lower first cost can be expected if the 480 wye/277-volt system is selected.

References

1. VOLTAGE UP—WIRING COST DOWN, A. I. Sawyer. *Electrical West*, New York, N. Y., Dec. 1952.

2. HIGH-VOLTAGE FLUORESCENT-LIGHTING POWER SYSTEMS FOR OFFICE BUILDINGS, R. R. Lang. *G. E. Technical Publication GET-2307A*, General Electric Company, Schenectady, N. Y.

3. NATIONAL ELECTRIC SAFETY CODE. National Bureau of Standards, Washington, D. C., 1953.

4. MODERN ELECTRIC - POWER - DISTRIBUTION IDEAS AS APPLIED IN A LARGE WAR PLANT, R. H. Kaufmann, N. A. Kieb. *AIEE Transactions (Electrical Engineering)*, vol. 64, June 1945, pp. 279–87.

◆

Discussion

Kurt-Beeman Paper

Robert F. Lawrence (Westinghouse Electric Corporation, East Pittsburgh, Pa.): Mr. Kurt and Mr. Beeman state, "With the autotransformer, there is a possibility of overvoltages on the secondary service through loss of the neutral or an open winding . . ." In the case of a 3-phase autotransformer of core-type construction, an open circuit in the winding between the neutral point and the low-voltage conductor would result in an insignificant overvoltage. Tests have been made which substantiate this statement and an analysis shows it to be true because of the inherent characteristics of the core-type construction and consequent flux paths.

No specific tests of an open-circuited common winding of a single-phase autotransformer come to mind. For this case also, there is a conductive connection from the high-side conductor through the autotransformer winding to the low-side conductor. The low-side conductor would have the same potential to neutral as the high-side conductor for an autotransformer with no load on the secondary. However, examination of the equivalent circuit of an autotransformer with load will show that, under open circuit of the common winding, the voltage across the load is equal to

$$\frac{Z_{\text{load}}}{Z_{\text{magnetizing}}} \times E_{\text{line to neutral of high-side conductor}}$$

$Z_{\text{magnetizing}}$ is the magnetizing impedance as measured from the high to low conductor with the common leg open-circuited and it has a high value. Thus, the voltage across the load for even a small load (high impedance) will usually be small. In other words, the load would hold the secondary to neutral voltage to approximately its normal value or less than its normal value.

Loss of a neutral or opening of a winding in the application of autotransformers under discussion is at worst an improbable occurrence. In the light of this discussion it is felt that the possibility of such an occurrence produces no more serious hazard than the ordinary use of electricity. Considerable successful experience has been gained in the use of autotransformers.

C. E. Asbury and **O. A. Lentz** (Commonwealth Associates, Inc., Jackson, Mich.): We agree with Mr. Kurt and Mr. Beeman that a major consideration in using large secondary network systems at higher voltages is the short-circuit duty at service entrances and associated distribution equipment. The problem has been presented and evaluated but there still appears to be a question as to whether the equipment is suitable for the high short-circuit currents at 480 volts or if the available short-circuit current should be limited to 100,000 amperes. It is realized that the short-circuit currents calculated by the general short-circuit methods may never be realized in practice. Even the small impedance of a network protector, leads, and contact resistor have a large effect in the magnitude of short-circuit current. It is believed that the current-limiting effect of these factors together with the burn-free feature on the 120/208-volt systems may have been major factors in their successful operation.

The same impedance factors are present in the 277/480-volt network but as an inverse function. It appears that contact resistance, arcing, etc., will have somewhat less effect on limiting the maximum short-circuit current.

Many technical papers have been presented on calculation of short-circuit currents on lower voltages. However, in discussing this problem there appears to be a difference in evaluating certain factors and as a result there are wide variations in the magnitude of short-circuit currents calculated. It would be helpful if there were one recognized method of calculation for use in applications of this nature.

When considering the application of protective equipment for the customer service, one either has to consider protective equipment having sufficient momentary and interrupting rating to withstand the high rated short circuits or limit the fault current by design of the network system. If fault currents are not limited it will be necessary that branch circuit as well as the main supply protective equipment be made available at a reasonable cost with high interrupting and momentary ability. The short-circuit current to the customer's bus may be limited by the design of the service entrance such as:

(1) dual feeds, (2) spacing the phase cables for high impedance, or (3) installing reactors at an additional cost. The fault currents can also be controlled to some extent by the selection of higher transformer impedance supplying the network providing voltage problems can be adequately worked out.

Referring to Table III of the paper, it is not clear in the cost comparison between the 208- and 480-volt system if the equipment cost for the 277/480-volt system is based on equipment design to withstand the high short-circuit current of 200,000 amperes or for a system which has been designed to limit the short circuit within the rating of present-day equipment. If the cost is on the basis of equipment suitable for 100,000 amperes, does the comparison include the cost for current-limiting equipment such as reactors? Assuming that the cost figures are based on lower rated equipment, it would be interesting to know what the protective equipment costs would be if it were adequate for the 200,000-ampere interrupting ability which the authors feel could be designed for.

Table I. Full Load Voltage, Light Load Voltage, and Voltage Spread, All in Per Cent, for the 208/120-Volt System with Varying Taps on the Autotransformers as Shown in Fig. 1

N_1 = Per-Cent Tap Above Normal	E_1 = Voltage at First Load		N_2 = Per-Cent Tap Above Normal	E_2 = Voltage at Last Load		Voltage Spread, Per Cent
	Full Load	Light Load		Full Load	Light Load	
100	98.988	99.750	100.0	96.597	99.131	3.153
100	99.011	99.778	100.5	96.960	99.525	2.813
100	99.034	99.797	101.0	97.321	99.916	2.595
100	99.057	99.821	101.5	97.680	100.306	2.626
100	99.079	99.844	102.0	98.038	100.694	2.666
100	99.102	99.868	102.5	98.394	101.081	2.687
100	99.124	99.891	103.0	98.748	101.466	2.728
100	99.147	99.915	103.5	99.100	101.849	2.749
100	99.169	99.938	104.0	99.451	102.231	3.172

Table II. Full Load Voltage, Light Load Voltage, and Voltage Spread in Per Cent for the 208/120-Volt System with Varying Autotransformer Taps as Shown in Fig. 2

N_1 = Per-Cent Tap Above Normal	E_1 = Voltage at First Load		N_2 = Per-Cent Tap Above Normal	E_2 = Voltage at Second Load		N_3 = Per-Cent Tap Above Normal	E_3 = Voltage at Last Load		Voltage Spread, Per Cent
	Full Load	Light Load		Full Load	Light Load		Full Load	Light Load	
100	99.054	99.766	100	97.635	99.400	100	96.275	99.048	3.491
100	99.052	99.765	100	97.671	99.447	101	97.055	99.892	2.837
100	99.051	99.764	100	97.705	99.492	102	97.830	100.730	3.025
100	99.107	99.820	101	98.470	100.286	101	97.127	99.977	3.159
100	99.106	99.819	101	98.506	100.333	102	97.904	100.818	2.914
100	99.161	99.875	102	99.301	101.167	102	97.976	100.903	3.191

Barnett-Lawrence Paper

D. M. Sauter (Westinghouse Electric Corporation, East Pittsburgh, Pa.): In the analysis of the allocation of voltage drops in a dual-voltage secondary system, the authors noted that the prospect of using different taps in the step-down autotransformers at different locations in the system should be investigated as a means of alleviating the problem of maintaining the proper voltage spread. This means that, if the autotransformers are served from a 277/480-volt radial and then are also interconnected with cable at the 120/208-volt level, an analysis of the performance of autotransformer taps cannot be approximated since the product of vector transformation ratios around the closed loops is not unity. Thus, with turn ratios other than unity, special consideration must be used in any analysis. Before presenting the general results obtained from analyses of simple systems, a review of the considerations involved in determining the voltage spread of a system appears worth while.

The voltage profiles illustrated in Fig. 3 of this paper show the voltage drops for light load and full load conditions for each of the two voltage levels of the secondary system. In determining the voltage spread of this system for one voltage level, two points were established on the profile: one at the minimum voltage of the profile and the other at the maximum voltage of the profile. The voltage difference between these two points is the voltage spread obtained for this system. Thus, it can be realized that the voltage spread is determined by the maximum difference in voltage between any two utilization locations in the system; this maximum difference can be established at nonsimultaneous times.

Another important point to note is that the voltage drops portioned to each part of the systems studied in this paper are not rigid. That is, other voltage drop apportionments also would be correct as long as the customer's utilization voltage remains within the preferred voltage range. This is an important feature since the variation of autotransformer taps can alter the direction and amount of current flow to the various loads.

In an attempt to study the effects of various values of taps on the transformers, systems similar to that illustrated in Fig. 3 of the paper were studied. The source transformers, cables, and autotransformers used in the various systems studied are illustrated in Figs. 1 and 2 of the discussion. The circuit then consisted of a spot network source with a secondary voltage of 480 volts. This source feeds two or three 500-kva autotransformers which are interconnected by 277/-480-volt cable as well as 120/208-volt cable. This cable consists of two 4/0 cables per phase except in Fig. 1 where the first span of 277/480-volt cable from the source has four 4/0 cables per phase since it is feeding two autotransformers. All the load was assumed at 120/208 volts and the loads were assumed to be a constant impedance which at full load would take 420 kva with a lagging power factor of 80 per cent when 100-per-cent voltage was applied. These loads then total 85 per cent of the installed autotransformer capacity.

Complex current nodal equations were then established for each load location and

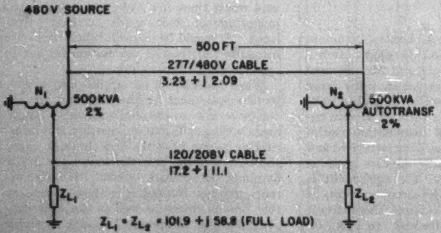


Fig. 1. Circuit diagram of autotransformer connections with one loop. All impedances in per cent on 500-kva base. N_1 and N_2 are tap positions of each autotransformer

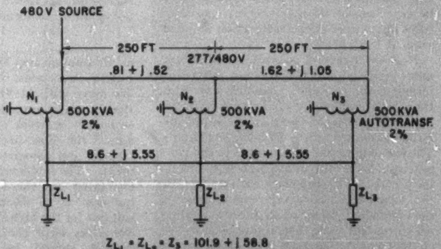


Fig. 2. Circuit diagram of autotransformer connections with two loops. All impedances in per cent on 500-kva base. N_1, N_2, and N_3 are tap positions of each autotransformer

these equations were solved simultaneously on the digital computer for the voltages at the various loads for a given autotransformer tap position. The results of these solutions are given in Tables I and II of the discussion, where the 120/208 voltages at the various load locations are tabulated for each system in per cent for full load and light load conditions. For light load conditions, the loads were assumed to be one-quarter of those existing at full load. Also shown in the last column of each table is the voltage spread, as defined earlier in this discussion, that exists in the system for the various tap combinations.

From the results of these studies, it was determined that the turns ratios effected a decrease in voltage spread of about 0.005 per cent. That is, under light load conditions, enough current is circulated around the loops by the off ratio taps to compensate for this much of the voltage rise caused by these higher tap positions. It should also be noted that these higher tap autotransformers have raised the voltage level of the extreme loads. Since there are many vagaries that exist on a given system in the nature of the loads and in the occurrence of the peak demand of these loads, taps may prove advantageous in certain circuit locations. However, it appears that each location must be examined on its own merits and trial installations made in the field. These field installations will establish a pattern to guide the further application of taps to reduce voltage spread and at the same time raise the voltage level of the system.

Joint Discussions

Robert A. Zimmerman and **Henry E. Lokay** (Westinghouse Electric Corporation, East Pittsburgh, Pa.): We believe that the conclusions of the Kurt-Beeman and Brereton-Donnelly papers are basically in accord with our paper.[1] Where there are differences of opinion, the differences are a matter of degree rather than of principle.

The cost-ratio points for the particular 3-floor building discussed in the Brereton-Donnelly paper fall within the bands of Figs. 1 and 2 of our paper. Since Figs. 1 and 2 deal with 1- and 12-floor buildings, this would indicate that the bandwidths are representative of practical cases of buildings in this size range.

The cost-ratio points from the Kurt-Beeman paper for the 10-, 20-, and 40-floor buildings fall just outside the bands of Figs. 2 and 3 of our paper. These curves are for 12- and 36-floor buildings. We feel that the reason that these points do not fall within the bands is that the final cost comparisons made in Tables I, II, and III of the Kurt-Beeman paper do not include all of the utilization devices. The large motors for the air conditioning have been included, but the other utilization devices such as the fluorescent luminaires, elevator motors, air-handling equipment, and miscellaneous motors have been omitted. As stated in the conclusions of reference 1, the utilization devices should all be included in any economic comparison of the total electric system costs since they are such a large part of the total cost.

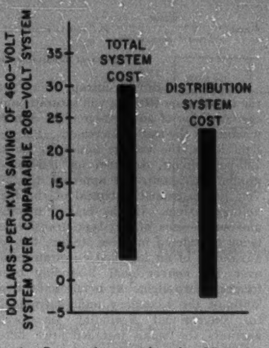


Fig. 1. Range of savings found in 35 investigations[1]

Fig. 1 of this discussion illustrates this point very well. The distribution system comparison shows that for some of the cases studied the 208-volt distribution system showed a saving over a comparable 460-volt distribution system. However, in all cases studied, the total cost comparisons favored the 460-volt system. The total cost comparisons included all the utilization devices.

In the 36-floor building investigations of reference 1, it was found to be impractical, because of voltage drop limitations, to use a 208-volt system with transformers located at basement level only. For this reason, the 208-volt studies had transformers on an upper floor as well as at basement level. This system was compared to a 460-volt system with transformers located at basement level only. To make a more parallel comparison between the two voltages, a 460-volt system was studied with transformers on an upper floor as well as at basement level. In the Kurt-Beeman paper, the transformers are located at basement level only for both 208 and 460 volts for the 40-floor building. We do not feel that this is a fair comparison inasmuch as it severely penalizes the 208-volt system. The results of this comparison favored the 460-volt system by $32.10 per kilovolt-ampere. This gives a ratio of 460-volt system cost to 208-volt system cost of 0.74. We do not feel that savings of this magnitude can be realized in a more practical case. In the 36-floor building study, it was found that the saving in favor of the 460-volt system was in the order of $15 per kilovolt-ampere.

The greatest saving that was found was for the 12-floor building with the transformers located at basement level for both voltages. This saving was in the order of magnitude of $30 per kilovolt-ampere in favor of the 460-volt system.

Reference

1. The Relative Feasibility of 460-Volt or 208-Volt Service in Commercial Buildings, H. G. Barnett, R. A. Zimmerman, H. E. Lokay. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 306–13.

A. J. Pansini (Long Island Lighting Company, Hicksville, N. Y.): Mr. Pansini's discussion is included with the discussions of the symposium papers which appear in the December issue of *Power Apparatus and Systems.*

W. R. Bullard (Ebasco Services, Inc., New York, N. Y.): This series of papers presents a wealth of information on a subject that is currently of very great interest to the electric power industry. The practice of serving large commercial buildings at voltages in the 440-volt class is expanding rapidly. The data on economics contained in some of these papers no doubt will add impetus to the movement, and the various other types of information will be helpful in the development of improved design and operating techniques.

One step of importance that must be taken relatively soon by the industry generally is that of selecting a specific design and operating voltage level for purposes of standardization. Studies of national scope with this in view are now in progress by the Edison Electric Institute and the manufacturers. The problem is a complex one, involving the entire field of electric appliances, and offering an opportunity for selecting a standard that will be adaptable to all needs in the 440-volt class for a great many years in the future. In view of these studies and the need for complete open-mindedness in weighing all factors, it is unfortunate that a tone of finality has crept into several of these papers when speaking of a specific voltage level in the 440-volt class, i.e., 265/460 volts (or 277/480 volts which in effect is a maximum rating for the same operating level). It is true that this level has gained some headway in actual application of the higher voltages in metropolitan areas, but this has occurred with relatively little thought to possible ultimate application to service of all classes and utilization devices of all kinds. Also the number of installations actually operating at this voltage level is microscopically small in comparison with probable future usage of service in the 440-volt class. It is to be hoped that the present small usage of the 265/460-volt level and the tenor of some of the papers in this series in speaking of this specific level will not unduly influence the thinking of the industry generally, pending the outcome of studies now in progress.

Closing Remarks

H. D. Kurt and **Donald Beeman**: Mr. Asbury and Mr. Lentz point up the problems of interrupting capacity on the 480-volt network system. All the cost comparisons considered with a secondary network service are based on an interrupting requirement at the entrance switchgear of 100,000 average 3-phase asymmetrical amperes. It has been assumed that the short-circuit current has been reduced by the use of high-impedance transformers and/or wider spacing of bus and/or reactors. The additional assumption has been made that all such equipment is located in the utility network vault and is therefore not a charge on the building service. If the short-circuit current at the building entrance is above 100,000 amperes, it is then proposed to use reactance in the form of wide spacing of bus or a reactor to limit this to 100,000 amperes. If a reactor is used, an approximate installed cost for a 4,000-ampere reactor of necessary impedance to reduce the short-circuit current from 200,000 amperes to 100,000 amperes is $5,500.

Mr. Zimmerman and Mr. Lokay mentioned the fact that the utilization devices such as fluorescent luminaires, elevator motors, air-handling equipment, and miscellaneous motors have been omitted from the cost comparison. It should be pointed out that all utilization equipment for which a price differential exists between the 208-volt and 480-volt system have been included in the cost comparisons. The smaller motors are the same price whether rated 208 volts or 480 volts. The only difference is in the control for such motors and the differential for such control has been included in the cost of the motor-control centers.

The fluorescent luminaires for operation at 277 volts may cost more, the same as or less than those used at 120 volts depending upon the type of ballast selected. Even for those fluorescent luminaires costing more at 277 volts, the price differential is small, and in large-quantity orders the differential can disappear. For this reason, the fluorescent luminaires have not been included in any of the cost comparisons.

Mr. Lawrence's comment on the stabilization of the neutral point with a 3-phase autotransformer of core-type construction is correct. Autotransformers of shell-type construction and also 3-phase units made up of three single-phase autotransformers do not exhibit the characteristic of stabilizing the neutral under the condition of loss of the neutral.

In the case of an open-circuited common winding of a single-phase autotransformer, the correct equation for the voltage across the load is

$$E_{\text{load}}=\frac{Z_{\text{load}}}{Z_{\text{load}}+Z_{\text{magnetizing}}}\times E_{\text{line to neutral of high-speed conductor winding}}$$

$Z_{\text{magnetizing}}$ is not a constant value since the transformer winding will saturate under the condition of an open-circuited common winding as the load is increased. For very light loads, the voltage appearing on the 120-volt circuit, designed for 150-volt to ground maximum, will approach the 277-volt line-to-ground potential of the high-voltage system. In the limited quantity and small sizes of dry-type transformers being considered in these studies, the cost advantage of an installed autotransformer over the conventional 2-winding transformer is very slight. It is felt that the isolation of the secondary fault currents by the conventional 2-winding transformer warrants the small installed price differential.

H. G. Barnett and **R. F. Lawrence:** We thank Mr. Bullard, Mr. Pansini, and Mr. Sauter for their discussions. As pointed out, EEI and NEMA committees are considering the problem of higher utilization voltages. Undoubtedly, the work that is being done will result in resolution of the many problems that are involved.

Since the papers constituting this symposium concentrate on the problems in commercial areas, we have covered tolerable voltages of the utilization equipments required in such areas. One possible exception is the 230-volt single-phase motor. Based on the data of Figs. 1 and 2 of our paper, the tolerable voltage of the 230-volt single-phase motor is unsuitable for line-to-line connection in the 120/208-volt system and for line-to-neutral connection in the 264/460-volt system. However, it is felt that the application of motors in commercial buildings are generally covered by the range of sizes in either the 115-volt single-phase or 220-volt 3-phase motors. The 230-volt single-phase motor is better suited to the 240/416-volt system.

Mr. Pansini's discussion points out the relation of 416 volts and 480 volts to 440-volt 3-phase motors. Fig. 2 of our paper indicates that the tolerable voltage for these motors is better suited to the 265/460-volt system, especially where motors are operated close to their rated output as would be expected in commercial buildings.

Mr. Sauter's analysis is very pertinent and effectively illustrates that the use of different taps in the step-down autotransformers can be beneficial. The discussion applies to the use of autotransformers in the spot-network radial system described in references 1 and 9 of the paper.

D. S. Brereton and **H. J. Donnelly:** The discussion by Mr. Zimmerman and Mr. Lokay substantiated, by their experience, the results of this comparative cost study. Realizing that care, particularly with small buildings, should be exercised in publicizing the results of such a study, it should be emphasized that an independent consulting engineer was obtained to prepare complete drawings and an electrical contractor obtained competitive bids on all electric devices and labor.

A High-Power Servo Analyzer

W. STERLING GORRILL
MEMBER AIEE

OMAR C. WALLEY
ASSOCIATE MEMBER AIEE

IN the design of a closed-loop control system, it is necessary to study the dynamic characteristics of the control circuit and its components to determine their effect upon accuracy, speed of response, and stability. While the advent of analogue computers has greatly simplified these studies, difficulties are encountered in setting up analogues that incorporate the same nonlinearities and operate at the same power levels as the equipment involved. To provide a means of obtaining more accurate information, the special high-power servo analyzer described in this paper was designed and constructed; see Fig. 1. Its function is to determine the frequency response curves, both magnitudes and phase shifts, of actual apparatus operating at normal power level. This servo analyzer has been used to study the characteristics of aircraft electric generating systems; however, it could be employed in the design of any type of control system. Thus, the development of the servo analyzer adds another design tool to those already available for control system design.

General Description

The servo test set has been designed to assist in the design of voltage regulation and frequency regulation equipment for rotating machines. Specifically, the test set aids in the solution of stability problems by measuring the transfer functions of either complete open-loop systems or of individual system components such as generators, alternators, inverters, carbon pile regulator assemblies, and damping transformers.

To determine the transfer function of a component, an input signal is applied to the component, the signal being sine-wave modulated at low frequency. The output signal of the component is ob-

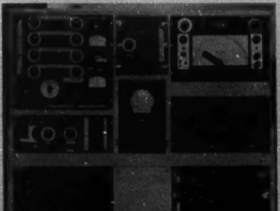


Fig. 1. Front view of servo analyzer

Paper 54-419, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 29, 1954; made available for printing September 13, 1954.

W. STERLING GORRILL is with Hanson-Gorrill-Brian, Inc., Glen Cove, N. Y., and OMAR C. WALLEY is with the Jack & Heintz Precision Industries, Inc., Cleveland, Ohio.

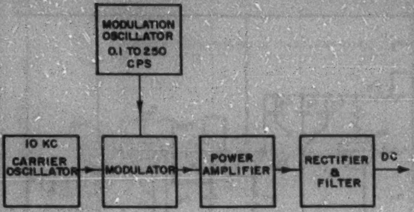

MODULATION OSCILLATOR 0.1 TO 250 CPS
10 KC CARRIER OSCILLATOR
MODULATOR
POWER AMPLIFIER
RECTIFIER & FILTER
DC

Fig. 2. Block diagram for modulated d-c signal

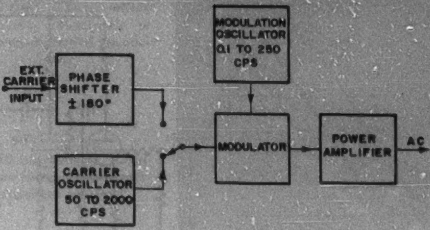

EXT. CARRIER INPUT
PHASE SHIFTER ±180°
MODULATION OSCILLATOR 0.1 TO 250 CPS
CARRIER OSCILLATOR 50 TO 2000 CPS
MODULATOR
POWER AMPLIFIER
AC

Fig. 4. Block diagram for amplitude-modulated a-c signal

served and, as the modulation frequency is increased, the attenuation and phase shift of the modulation is measured. Complete open-loop systems can be studied in the same manner.

In considering the components and systems which might be studied, it was clear that an unusually versatile test set was desired and that conventional servo analyzers were entirely inadequate. For example, suppressed carrier signals, while common in position servo systems, are seldom met in voltage or frequency regulators. On the other hand, modulation frequencies as high as 250 cycles per second (cps) may be important and frequency-modulated signals are common. Modulated input signals of large power are desired so that the field of a moderate sized generator, for example, can be excited directly. Phase shift and attenuation measurements must be possible from very small carrier voltages so that low-resistance shunts can be used where output current modulation is the desired variable.

The measuring equipment is functionally separate from the signal generating equipment with the exception of the modulation phase reference required in determining phase shift in the components under test. The two groups of equipment are described in the following.

Signal Generating Equipment

Purely electronic means of signal generation are used because of the very wide ranges of both modulation and carrier frequencies involved. Fig. 2 is a simplified block diagram of the equipment used in generating modulated direct voltages. It is seen that a 10-kc carrier is modulated, amplified, and then rectified and filtered to give the desired modulated direct voltage. This method was chosen because it provides the desired wide range of output currents and voltages by the relatively simple expedient of switching output transformer taps and rectifiers. The modulator and power amplifier are thus used for both a-c and d-c carriers.

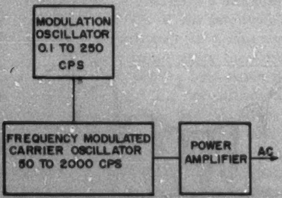

MODULATION OSCILLATOR 0.1 TO 250 CPS
FREQUENCY MODULATED CARRIER OSCILLATOR 50 TO 2000 CPS
POWER AMPLIFIER
AC

Fig. 3. Block diagram for frequency-modulated a-c signal

The 10-kc carrier is generated by a simple fixed frequency oscillator and modulated in a conventional push-pull modulator using two variable mu tubes with transformer input and output circuits. The modulation signal is generated by a low-frequency oscillator and applied to the center tap of the input transformer secondary. Fig. 2 shows the modulation frequency range actually used, 0.1 to 250 cps, rather than the full range of the oscillator.

The commercially available oscillator used is uniquely suitable for the application because square waves, symmetrical saw tooth, and sine waves are generated simultaneously. Normally only one wave shape is available at a time. A simple modification has made possible the simultaneous use of sine waves for modulation and saw tooth for phase reference signal to the measuring equipment. The square-wave signal is also available when desired for transient study. The power amplifier was designed to meet the following special conditions:

1. Lowest possible output impedance to minimize source impedance effects on components under test.

2. Normal use into mismatched loads. Maximum voltage may frequently be required, but maximum power will only be used occasionally.

3. Adequate frequency response, low distortion, and power output. Efficiency, weight, and space are not important.

To meet these conditions a pair of *304 TL* tubes are used with class *A* operation giving a power output of 180 watts into a resistive load. Negative feedback of 20 decibels (db) is used from the primary of the output transformer to reduce output impedance. Feedback from the secondary would be desirable but is impossible

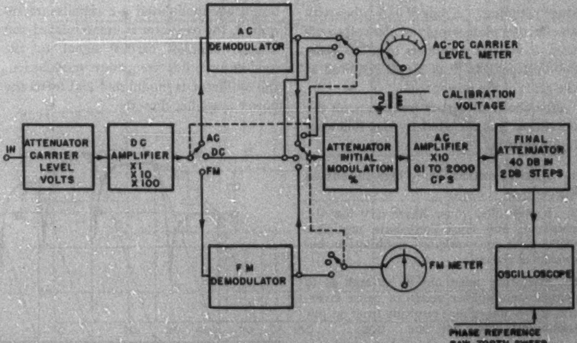

AC DEMODULATOR
AC-DC CARRIER LEVEL METER
CALIBRATION VOLTAGE
IN
ATTENUATOR CARRIER LEVEL VOLTS
DC AMPLIFIER X1 X10 X100
AC
DC
FM
ATTENUATOR INITIAL MODULATION %
AC AMPLIFIER X10 0.1 TO 2000 CPS
FINAL ATTENUATOR 40 DB IN 2DB STEPS
FM DEMODULATOR
FM METER
OSCILLOSCOPE
PHASE REFERENCE SAW TOOTH SWEEP

Fig. 5. Block diagram of measuring circuits

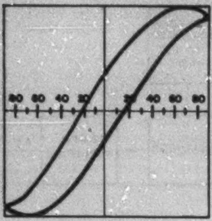


Fig. 6 (above). Oscilloscope pattern for modulated d-c signal with 20 degrees' phase shift in equipment being studied

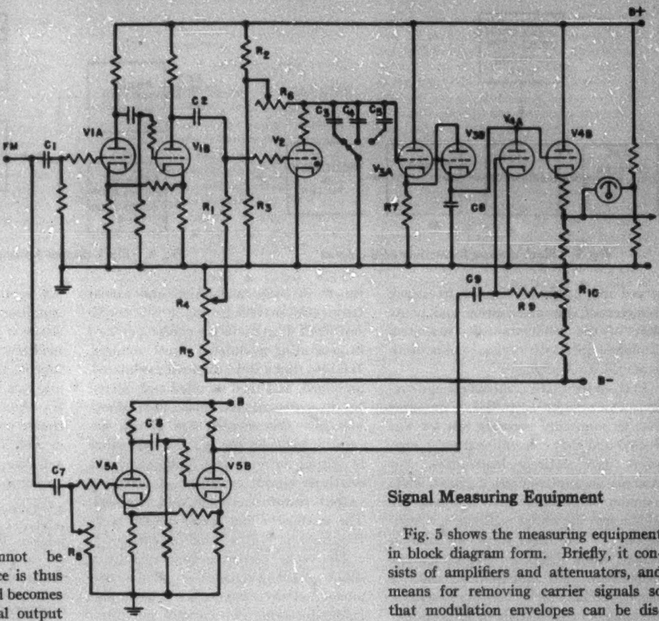


Fig. 7 (right). Circuit diagram of frequency-modulated demodulator

because the secondary cannot be grounded. Secondary resistance is thus not reduced by the feedback and becomes an important part of the total output impedance. The problem of providing the correct voltage taps for both alternating and direct voltages and still maintaining low secondary resistance led to the use of two limited-range output transformers: one for 50 to 2,000 cycles, and the second for 10-kc only. The 1,850-volt plate supply is automatically removed before switching transformers to avoid excessive switching transients.

Stability with 20-db feedback was not a difficult problem as the class-*A* operation allowed resistance-capacitance interstage coupling. A pair of *6L6* tubes with an effective plate supply of 550 volts were used as voltage amplifiers to provide the 290-volt peak-to-peak signal required at the grids of the power tubes.

The choice of carrier frequency for d-c operation was dictated by the following requirements:

1. Efficient rectification of high currents at low voltage.
2. Ripple filter must have very low d-c resistance and must contribute negligible phase shift at maximum modulation frequency of 250 cps.
3. Frequency must not be so high as to complicate amplifier design or cause excessive source impedance resulting from output transformer losses.

These conditions were met by the use of full-wave germanium power rectifiers operating at 10 kc. Figs. 3 and 4 are block diagrams of signal generating equipment for amplitude and frequency-modulated a-c signals. The power amplifier and modulation oscillator are the same units as used for d-c signals. For carrier signals, a standard sweep-frequency oscillator was modified to receive external modulation signals and to give stable sine-wave frequency modulation of the low amplitudes desired. When amplitude-modulated a-c signals are required, the oscillator is unmodulated and simply supplies carrier signal to the modulator. For frequency modulation, the oscillator is modulated and feeds the power amplifier directly.

Signal Measuring Equipment

Fig. 5 shows the measuring equipment in block diagram form. Briefly, it consists of amplifiers and attenuators, and means for removing carrier signals so that modulation envelopes can be displayed on the cathode-ray oscilloscope screen. Carrier level is measured by adjusting the input attenuator for a fixed reading, 20 volts, of the carrier level meter and then reading the voltage on the calibrated attenuator dial. Modulation amplitudes are then read by adjusting the scope pattern to a predetermined size and reading attenuator dials.

The symmetrical saw-tooth voltage from the modulation oscillator is connected to the horizontal amplifier of the cathode-ray oscilloscope while the modulation-envelope voltage drives the vertical amplifier. Since the saw-tooth voltage rises and falls linearly with time, the intercept of the Lissajous pattern with the linearly calibrated horizontal axis of the cathode-ray oscilloscope screen gives

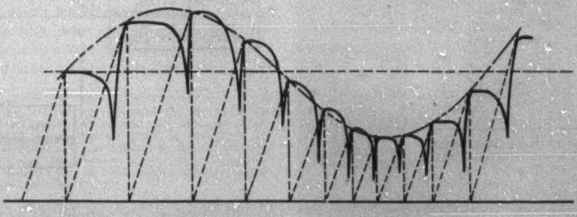


Fig. 8. Frequency-modulated demodulator wave forms

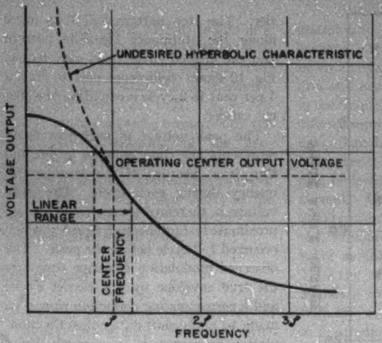


Fig. 9 (left). Output voltage—frequency characteristic of frequency-modulated demodulator

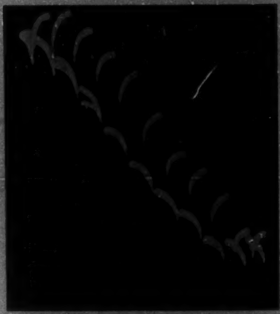


Fig. 10 (upper right). Oscilloscope pattern for frequency-modulated a-c signal

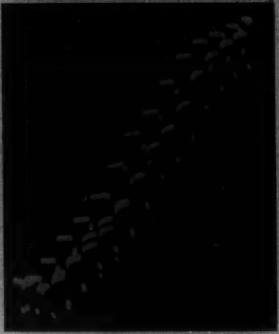


Fig. 11 (lower right). Oscilloscope pattern for amplitude-modulated a-c signal

modulation phase shift of the component under test. Fig. 6 is the pattern observed for a 20-degree phase shift shown against the screen calibrated in degrees of phase shift. The use of the cathode-ray oscilloscope for all signal presentation has the advantage of providing constant monitoring of the modulation wave form. It signals the presence of nonlinear conditions where transfer function information may be meaningless.

An even more important reason for cathode-ray oscilloscope presentation is that it greatly simplifies the problem of demodulating a-c and frequency-modulated signals without introducing undesired phase shifts. Fig. 7 is a schematic of the frequency-modulated demodulator. V_{1a} and V_{1b} amplify and make square waves of the incoming frequency-modulated signal. C_2 and R_1 differentiate the square wave so that spikes are formed to fire the *884* gas tube (V_2) at the beginning of each cycle. V_2 together with R_6, and capacitor C_3 form a typical saw-tooth generator which, however, is not free-running since the grid of V_2 is negative except when triggered from V_1.

With any fixed setting of R_6, and temporarily assuming linear charging of C_3, the amplitude of the saw tooth will be proportional to the time between successive cycles and is, thus, a measure of frequency. In Fig. 8, the saw tooth is shown in dotted lines for a sine-wave frequency modulation where the carrier is roughly 10 times the modulation frequency. V_{3A} and V_{4B} are cathode followers and V_{3B}, C_8, and V_{4A} form a "pulse-stretching" circuit to give a stepped approximation to the desired modulation envelope. C_8 is charged through diode V_{3B} to the peak voltage of each saw tooth. Since V_{4A} is normally cut off by a strong negative bias, the peak voltage is maintained across C_8 during the following cycle. A short positive pulse from squaring amplifier V_5 and differentiating network C_9 and R_9 makes V_{4A} conducting just before the next saw-tooth peak, and discharges C_8 to ground. C_8 is then immediately recharged to the peak value of the next saw tooth. C_7 and R_8 are used to advance the phase of the signal entering V_5 so that the discharging takes place just before the saw tooth peaks.

If the saw-tooth rise were perfectly linear, the level of each output step would be proportional to the time between cycles and inversely proportional to frequency. The resulting hyperbolic output characteristic would distort the modulation envelope. To avoid this, capacitor C_3 is allowed to charge to a voltage where the nonlinearity of its exponential charging characteristic compensates for the hyperbolic nonlinearity. As shown in Fig. 9, a fairly linear voltage frequency characteristic is achieved up to roughly 15 per-cent modulation. A

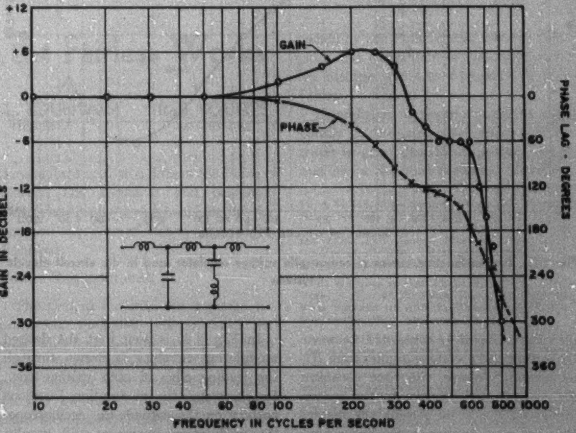


Fig. 12. Transfer-function curves of low-pass filter

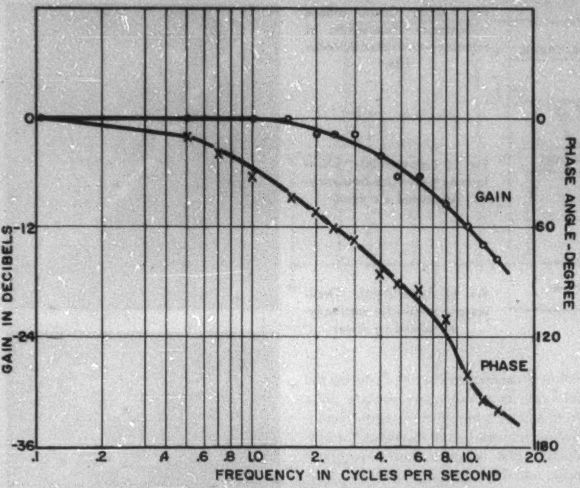


Fig. 13. Transfer-function curves of exciter and alternator for aircraft

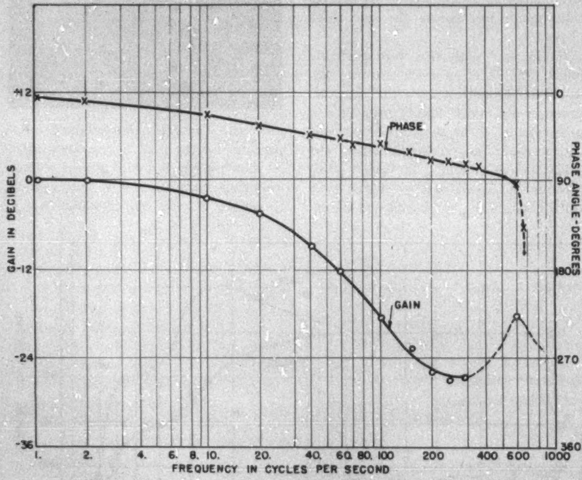


Fig. 14. Transfer-function curves of carbon-pile voltage regulator used in d-c aircraft electric systems

zero-center meter is connected between the output and a voltage divider from B^+ to ground. In use, the time constant R_8C_8 is adjusted relative to the center carrier frequency by zeroing the meter. This insures operation in the linear portion of the characteristic.

In Fig. 8 it is seen that the desired modulation envelope is drawn through the leading edge of each voltage step. In observing the modulation Lissajous pattern on the cathode-ray oscilloscope, the modulation frequency will usually not be an integral submultiple of the carrier. The step pattern will thus move along the stationary Lissajous pattern and the desired envelope is easy to see. Fig. 10 shows oscilloscope pattern from a 4-per-cent 18.2-cycle modulation of a 400-cps carrier.

The peak voltage of each saw tooth represents the time between cycles and thus is a measure of the average frequency during each cycle. The peak voltage is, therefore, a measure of the approximate instantaneous frequency which occurred 1/2 cycle before the peak. The observed modulation envelope thus lags the true envelope by 1/2 carrier cycle; and a corresponding correction should be made to phase-shift data unless the modulation frequency is very much lower than the carrier frequency.

The a-c demodulator shown in Fig. 5 is similar to the frequency-modulated demodulator in that it uses a pulse-stretching circuit to form a stepped modulation envelope. This avoids the use of filters and undesired phase shift which would be required if a simple rectifier detector were used. Fig. 11 shows the stepped modulation envelope from a 400-cycle carrier modulated at 13.3 cps.

Again referring to Fig. 5, the output of either the a-c or the frequency-modulated demodulator is a modulated d-c signal. The a-c amplifier eliminates the d-c component and provides necessary additional gain. This amplifier must pass modulation frequencies from 0.1 to 250 cps, but must also pass much higher frequencies so as not to spoil the step pattern from the demodulators. It consists of a stable d-c amplifier with the input signal coupled through a very large capacitor to block the d-c component.

When measuring the transfer function of a component, the first step is to apply a reasonable large modulated signal to the component at a frequency well below the point where phase shift and attenuation is expected. The output of the component is connected to the measuring equipment and the carrier-level and initial-modulation attenuators are set so that the Lissajous pattern just fills the square of the cathode-ray oscilloscope screen. As the modulation frequency is increased, the final attenuator is adjusted to maintain the oscilloscope pattern at constant amplitude. This attenuator is calibrated in 2-db steps and provides data directly in db-attenuation versus frequency.

Four input channels are provided in the measuring equipment. One channel is permanently connected to monitor the output of the signal-generating equipment. The remaining three channels may be connected to various points in a system under test so that the attenuations and phase shifts of the several components in the system can be measured with a minimum of work. A switch selects the channel to be presented on the oscilloscope. Monitoring of the generated signal allows measurements even beyond the nominal modulation range of 0.1 to 250 cps since phase shifts in the test set itself can be observed and corrections made to final data.

Applications

Transfer functions of a large variety of equipment have been determined with the servo analyzer. Fig. 12 shows frequency-response and phase-shift curves of a 400-cycle low-pass filter. Fig. 13 gives transfer-function curves for an alternator whose field is powered by an integrally mounted exciter. The two field time constants result in typical curves which approach 12 db for octave attenuation slope and 180 degrees total phase shift respectively. Fig. 14 is the transfer function of a carbon-pile voltage regulator. The primary 6-db attenuation is caused by inductance of the magnet coil. The sharp peak at 600 cycles is caused by a mechanical resonance.

Summary

The servo analyzer has been built specifically to measure component and system transfer functions in voltage regulation and frequency regulation servo systems. It differs from conventional servo analyzers in the types of modulation and carrier signals, power level, and in the frequency ranges required. Modulation frequencies range from 0.1 to 250 cps. Signals may be amplitude-modulated direct voltage on alternating voltage, or frequency-modulated alternating current. Carrier frequencies from 50 to 2,000 cps can be used. Significant phase-shift measurements are possible with sine-wave modulation frequencies as high as 1/10 the carrier frequency. Modulated signals of up to 180 watts can be generated with low source impedance. Phase-shift and attenuation measurements can be made with carrier amplitudes as low as 50 millivolts.

Appendix. Servo-Analyzer Specifications

Signal Generating Equipment

The modulation may be of sine-wave or square-wave type with amplitude modulation for direct and alternating voltages and frequency modulation for alternating voltages. The modulation frequency may be from 0.1 to 250 cps. The equipment delivers up to 40-per-cent amplitude modulation and up to 40-per-cent frequency modulation. The carrier may be direct or alternating voltage in the range from 50 to 2,000 cps. The output power and source impedance are as follows:

0 to 10 volts alternating and direct: 5 amperes. Impedance 0.5 ohm for alternating voltages, 1.0 ohm for direct voltages.

0 to 40 volts alternating and direct: 3 amperes. Impedance 1.5 ohms for alternating voltages, 4.0 ohms for direct voltages.

0 to 120 volts alternating and direct: 1 ampere. Impedance 3.5 ohms for alternating voltages, 20.0 ohms for direct voltages.

Where phase of carrier is critical, external carrier can be used. A phase shifter is included for 50 to 2,000 cycles, 360 degrees' shift.

Measuring Equipment

Modulation may be amplitude-modulated alternating voltage or direct voltage, or frequency-modulated alternating voltage. Two types of measurements are used: 1. Relative amplitude of modulation in 2-db switch steps. Estimate to less than 1 db depending upon signal-to-noise ratio. 2. Relative phase shift to approximately 5 degrees. Accuracy depends upon signal-to-noise ratio and ratio of carrier to modulation frequency. Approximate measurements are made of carrier amplitudes (volts d-c or rms) and of modulation (per cent).

Carrier voltages may be between 50 millivolts and 200 volts. Measurements are normally possible from less than 1-per-cent modulation up to suppressed carrier conditions for direct voltages, up to 95-per-cent modulation for alternating voltages, and up to 15-per-cent modulation for frequency-modulated signals. Limits and accuracies depend upon signal-to-noise ratio and ratio of carrier to modulation frequency.

All amplitude measurements are made by reading the attenuator settings for fixed output. Phase measurements are made by means of a linear calibrated oscilloscope screen.

◆

No Discussion

A-C Power System at the Fairless Works

S. S. WATKINS
FELLOW AIEE

W. A. DERR
MEMBER AIEE

L. L. FOUNTAIN
FELLOW AIEE

R. B. SQUIRES
MEMBER AIEE

THE Fairless Works of the United States Steel Corporation, located on a bend in the Delaware River near Morrisville, Pa., was designed as a completely integrated plant, capable of processing raw materials into finished steel products, ready for industrial use. Seldom in the history of steel mills have engineers had the opportunity to design so complete an undertaking in its entirety, and thereby to take advantage of past experience from both technical and economic standpoints. This paper deals with the electric power system which supplies the loads of the various mill areas. Fig. 1 is a scale drawing which shows the high-voltage substations and transmission lines of the Fairless Works. The technical and economic considerations used as a basis for the design and operation of this system will be discussed.

Power Requirements

The present facilities are designed for producing 1,800,000 tons of steel per year, with land and the arrangements to permit future expansion. The electric power system was designed to supply 115,000 kw total. The electric power plant includes two 30,000-kw 13.8-kv 60-cycle generators. The additional power, to make up the total of 115,000 kw, is furnished by the Philadelphia Electric Company and the system of the Public Service Electric and Gas Company of New Jersey.

In addition to electric power, a completely co-ordinated steel works requires water in great quantities, process steam, and air for the blast furnaces. The Fair-

---

Paper 54-348, recommended by the AIEE Industrial Power Systems and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing July 22, 1954.

S. S. WATKINS is with Gibbs & Hill, Inc., New York, N. Y.; W. A. DERR, L. L. FOUNTAIN, and R. B. SQUIRES are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

A great deal of the credit for the successful completion of this project should go to the United States Steel Corporation engineers, who carried out the work of co-ordinating the various contracts with the numerous manufacturers who supplied the electrical equipment. H. G. Frostick, R. W. Sheffler and J. F. Headlee of the United States Steel Corporation Engineering Organization are to be particularly commended for their work on the over-all project, as is H. S. Eubank of Gibbs and Hill, Inc., for his design work on the powerhouse.

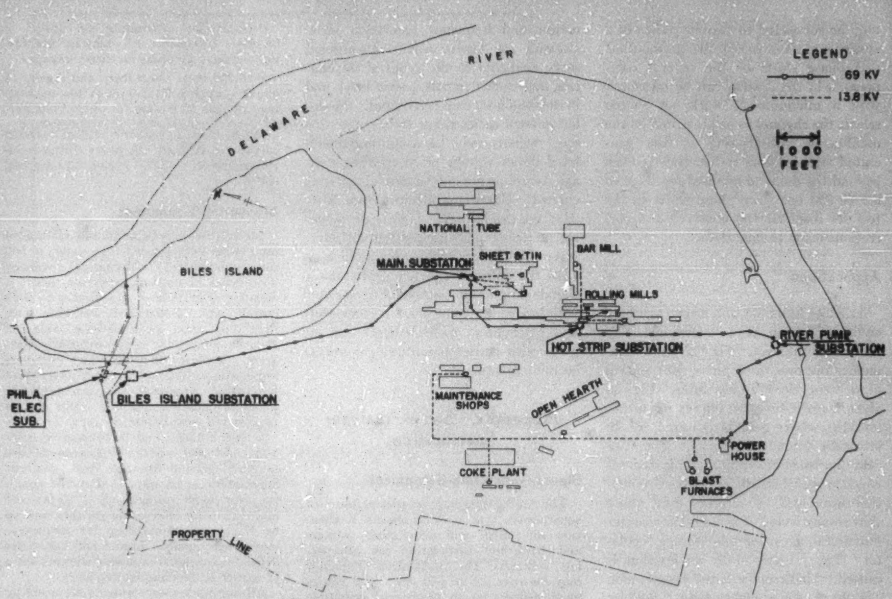


Fig. 1. Scale drawing showing transmission lines and high-voltage substations

less Works power plant includes three turbine-driven blowers for supplying air to the blast furnaces. Gases from the blast furnaces are burned in the boilers of the power house, which combines the functions of blower house and electric generating station. Provision is made for burning coke-oven gas, diluted natural gas, or fuel oil, whenever sufficient blast-furnace gas is not available. The design of the power plant is not within the scope of this paper, except as it affects the design of the electric power system.

Study of Power Transmission Systems

The fact that this was to be a completely new works made it possible to design the electric system for minimum costs and maximum flexibility. Six different types of power transmission systems were studied. These differed in voltage level, and consisted of either radial or loop feeds with reactor components to reduce fault currents where needed. The a-c network calculator was utilized in making these studies, which made it possible to check fault currents, voltage regulation, and other pertinent characteristics in the operation of a system. Fig. 2 gives the basic single-line diagrams of the various systems studied. It was found that all of these could be operated satisfactorily from an electrical standpoint, but an economic analysis showed that their costs differed widely, as summarized in Table I. The large amount of power involved made the cost of the 13.8-kv cable systems prohibitive, compared to distribution over aerial lines at 69 kv. In addition, at the time the plant was built, copper was in critically short supply so that the use of the higher system voltage made it possible to conserve this strategic material, and gave a reliable, flexible system which could be expanded in the future at minimum cost.

Table I. Comparison of System Designs Considered

Arrangement	Workability	Growth Possibility	Reliability	Cost, Per Cent
A—13.8 kv-loop cable, feeder reactors	excellent	fair	excellent	192
B—13.8-kv radial cable, feeder reactors	good	fair	good	137
C—13.8-kv radial cable, bus-tie reactors	good	fair	good	124
D—132-kv to 2 substations	good	good	good	114
E—132-kv to all substations	good	excellent	good	113
F—69-kv to all substations	good	excellent	good	100

Power System Installed

The power transmission system chosen was similar to that shown in Fig. 2(F), with some modifications dictated by later requirements. A single-line diagram of the system as installed and as shown in Fig. 3, indicates the 69-kv breakers used at all substations, to increase the flexibility while maintaining considerable economic advantage over any of the other schemes studied.

A double-circuit 69-kv overhead transmission line runs from the Power House 13.8/69-kv Substation to the Biles Island Substation, where power at 132 kv is received from the Philadelphia Electric Company. The two 69-kv transmission circuits pass through three stepdown substations as follows:

1. The Main 69-kv Substation steps down

to 13.8 kv for supply to the Cold Mills Substations.

2. The Hot Strip Mill 69-kv Substation steps down to 13.8 kv for supply to the Hot Mills Substations.

3. The Pump House 69-kv Substation steps down to 2.4 kv for supply to the river pumps and sewage and waste disposal equipment.

Underground 13.8-kv feeders from the Power House supply the electric power required by the maintenance shops, by-product coke ovens, blast furnaces, and open-hearth furnaces.

Biles Island Substation

In the Biles Island Substation the two 69-kv circuits of Fairless Works are connected through two 132/69-kv transformers to the 132-kv lines of the Philadelphia Electric Company. One of the lines goes directly to the Emilie Substation, while the other one taps the tie-line between the Emilie Substation and the Public Service Electric and Gas Company at Trenton, N.J. The two transformers at the Biles Island Substation are triple-rated 30/37.5/50 megavolt-amperes (mva) with the rating dependent on the method of cooling used. The tap-changing under-load equipment can be controlled either by the voltage at the Biles Island Substation or by remote control from the Power House.

The transformers at Biles Island Substation do not have individual circuit breakers on the 69-kv side, but there is a tie breaker between the two 69-kv lines. In the event that one of the 69-kv transmission lines is out of service, both transformers may be used in parallel over the remaining 69-kv line by closing the tie breaker and opening the line-disconnecting switches to isolate the line that is out of service.

Main and Hot Strip Mill Substations

At the Main Substation both the 69-kv and 13.8-kv circuits are outdoor circuits as shown in Fig. 4. At the Hot Strip Mill Substation the 69-kv circuits are outdoor as shown in Fig. 5, and the 13.8-kv circuits indoor as shown in Fig. 6.

Because the electrical arrangements at both substations are similar, they will be described together by reference to Figs. 3 and 7. Each of the 69-kv transmission lines is normally connected to its own 69-kv to 13.8-kv transformer. There are circuit breakers in each of the individual line sections coming into the substation and normally closed motor-operated disconnecting switches which connect the

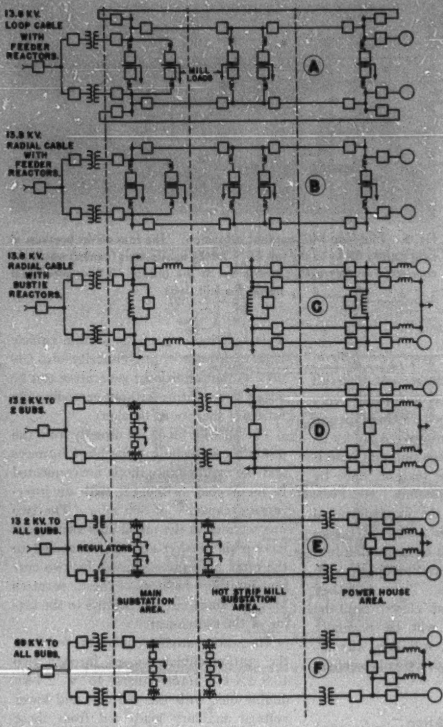
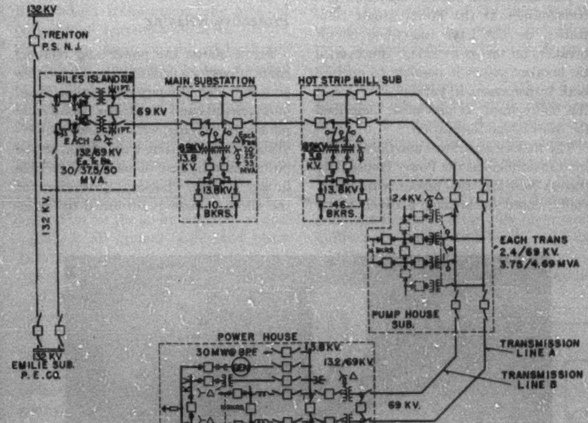

13.8 KV. LOOP CABLE WITH FEEDER REACTORS
MILL LOADS
A
13.8 KV. RADIAL CABLE WITH FEEDER REACTORS
B
13.8 KV. RADIAL CABLE WITH BUSTIE REACTORS
C
13.2 KV. TO 2 SUBS.
D
13.2 KV. TO ALL SUBS.
REGULATORS
E
MAIN SUBSTATION AREA.
HOT STRIP MILL SUBSTATION AREA.
POWER HOUSE AREA.
69 KV. TO ALL SUBS.
F

Fig. 2. Power transmission systems analyzed in determining the best system for the Fairless Works

Fig. 3 (below). Primary power system at the Fairless Works


132 KV
TRENTON P.S. N.J.
BILES ISLAND SUB
W/PT.
69 KV
EACH
132/69 KV
30/37.5/50 MVA.
132 K.V.
MAIN SUBSTATION
69 KV
13.8 KV.
Each Trans. 20 25 33 MVA
13.8 KV
10 BKRS.
HOT STRIP MILL SUB
69 KV
13.8 KV.
13.8 KV
46 BKRS.
2.4 KV.
BKRS.
EACH TRANS 2.4/69 KV. 3.75/4.69 MVA
PUMP HOUSE SUB.
132 KV
EMILIE SUB. P. E. CO.
POWER HOUSE
30 MW @ .8 P.F.
13.8 KV.
132/69 KV
GEN
69 KV.
TRANSMISSION LINE A
TRANSMISSION LINE B
Each Trans. 20-25-33 MVA.
30 MW @ .85 P.F.

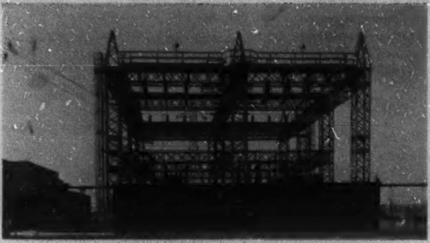


Fig. 4. Main outdoor substation. Structure for 69-kv circuits and the two 69/13.8-kv transformers in background. Throat-connected outdoor metal-clad switchgear in foregound. The duplex control switchboard is in the house at the left

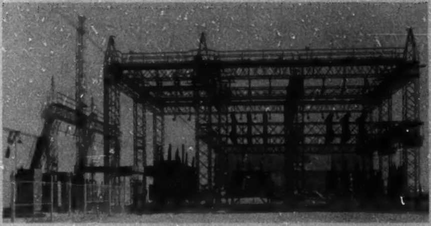


Fig. 5. Hot Strip Mill outdoor substation. The four 69-kv breakers at right supply 69 kv to the two 69/13.8-kv transformers (center) which are connected by segregated phase bus to the indoor 13.8-kv switchgear inside the mill (left)

transformer to its 69-kv bus. Each transformer is connected on the low-voltage side through a 13.8-kv circuit breaker to its individual 13.8-kv bus. A normally closed 13.8-kv bus tie circuit breaker is provided.

Under normal operating conditions, with all 69-kv breakers closed, the substation can be fed over four different transmission line sections. The tie breakers at each end of the 69-kv system permit power to be fed completely around the loop even though one line section or a power source is out of service. In the event that both sections of one of the 69-kv lines are out of service at one of the substations, full capacity is made available by closing the normally open 69-kv tie-disconnecting switch. All that is necessary to restore full capacity to the substation is to open the motor-operated disconnecting switch on the dead line and to close the motor-operated tie-disconnecting switch, so that both transformers are fed from the same line.

Individual mill loads are fed from both sides of the 13.8-kv substation bus by two 13.8-kv cable circuits. At the mill, the two cable circuits terminate on a single bus. This provides two sources of feed from the substation, which in turn has four routes of power flow normally available.

Pump House Substation

Because of the extremely critical nature of the cooling-water supply to the blast furnaces and other parts of the steel mill, a completely separate substation is devoted to the river pumps, even though the total load is relatively small. At present, this substation is equipped with two 69/2.4-kv transformers double-rated at 3,750–4,687 kva. It is contemplated that a duplicate set of transformers will be installed in the future. The present installation has an arrangement of circuit breakers and disconnecting switches very similar to that described for theMain and Hot Strip Mill Substations. A 2.4-kv tie breaker is used, but each pumping load is fed from only one bus section. The tie is normally closed at present. However, in the future, when the two additional transformers and their 2,400-volt busses are installed, they will be connected into the existing switchgear line-up through two more tie breakers. These will then normally be operated closed but the original tie breaker will be normally operated open. This will keep the interrupting duty on the 2,400-volt breakers within their rating.

Power House Substation

Just as at Biles Island Substation, the transformers at the Power House Substation do not have individual circuit breakers on the 69-kv side. However, a tie breaker makes it possible to operate both transformers in parallel over one of the 69-kv lines, as previously described for Biles Island Substation.

The 13.8-kv generators and the 69/13.8-kv transformers at the Power House Substation feed into two busses with a normally closed tie breaker between them. This switchgear is of the station cubicle type with interrupting capacity of 1,500 mva so that additional generation can be added in the future without exceeding the rating of the circuit breakers.

The 13.8-kv loads fed directly from the powerhouse, including the blast furnaces and the maintenance shops, are connected to metal-clad switchgear with an interrupting capacity of 500 mva. The two bus sections in the metal-clad switchgear have a tie breaker and are connected to the main powerhouse busses by two current-limiting reactors. These reactors limit the short-circuit currents to the rating of the switchgear.

The station auxiliary loads are fed from the station cubicle busses through two 13.8/2.4-kv transformers to a 2.4-kv double bus, with the 480-volt and lower voltage auxiliary loads fed from these busses through step-down transformers.

Protective Relaying

Fig. 7 shows the power and relaying connections for a typical transmission substation. The primary protection of all 69-kv circuits and all 13.8-kv cables is by pilot-wire relaying, with pilot-wire supervision relays to sound an alarm at the powerhouse in case of pilot-wire trouble in any section. Backup protection for the 69-kv circuits is furnished by direc-

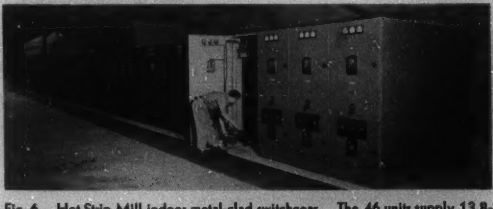


Fig. 6. Hot Strip Mill indoor metal-clad switchgear. The 46 units supply 13.8-kv power to the hot mill loads

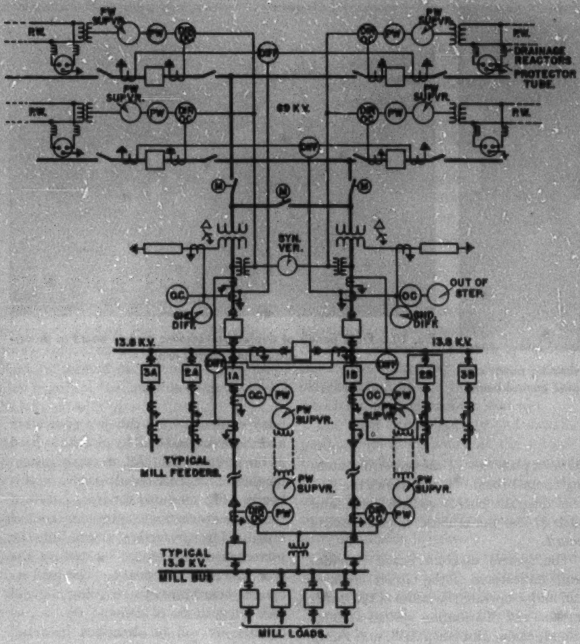


Fig. 7. Power and relaying connections at a typical 69- to 13.8-kv transmission substation

tional overcurrent relays. Because the 69-kv system is solidly grounded at the Biles Island transformers, the level of ground-fault current approaches that of the 3-phase fault current, and permits ground-fault protection to be furnished by the phase relays. This fact also permits the elimination of all 69-kv potential transformers, the directional overcurrent relays being polarized from potential transformers on busses at each substation. A network calculator study and a relay study showed that this backup relaying system for the 69-kv circuits was adequate for all ground faults and phase faults even with the maximum load current flowing in any line section

On the 69-kv system, it is necessary that each of the parallel lines be of sufficient capacity to carry the total load in the event that the other line is out of service. This means that the current transformer ratios are high in comparison to the normal full-load current. Because the transformers at Biles Island Substation can be fed on the 132-kv side from a bus tie and a line circuit breaker, it is necessary to use a 3-restraint transformer differential relay. To secure more sensitive protection for internal ground faults on the 69-kv windings of the Biles Island transformers, a high-speed induction type of ground differential relay is provided.

The ground current at all the 13.8-kv substation busses is limited to 1,600 amperes by a resistor in the power transformer neutrals. These transformers are also protected by a 3-restraint over-all transformer differential relay, plus a high-speed induction-type ground-differential relay, for increased sensitivity on ground faults near the neutral in the 13.8-kv windings of the transformers.

A transient stability study made on the electric system of the Fairless Works when connected to the utility system showed that the high-speed relaying makes it stable for all 69-kv or 13.8-kv faults. However, the possibility exists that the powerhouse generators could be tied into the utility system through two of the transformers in series at the step-down substations. This could come about, for example, if one 69-kv line *A* section was down for maintenance and a fault occurred on the *B* line on the other side of the substation. This would cause all synchronizing power to flow from the remaining 69-kv line down to the 13.8-kv bus through the transformer and to return to the other 69-kv line through the second transformer. These studies showed that under this condition the system might pull apart, especially if the line outages occurred on either side of the Pump House Substation with its relatively small-capacity transformers. Under these conditions it was important that the systems be separated quickly to maintain continuity of service on each half.

Out-of-step tripping relays were used at each of the three 69-kv substations to recognize this condition and to separate the system on the first swing if the angle between the powerhouse generators and the utility system approached 180 degrees. Separation is achieved by tripping of the 13.8-kv transformer breakers.

13.8-Kv Feeder Circuit Relaying

A typical 13.8-kv feeder circuit is shown in Fig. 7 from the 13.8-kv bus of the transmission substation to the individual mill bus. The primary protection is by pilot-wire relaying with backup time-overcurrent relays at the transmission substation and directional overcurrent relays at the mill substation end. The supervision relays on the pilot wires, in addition to indicating short circuits, grounds, and opens, are also used to transfer trip in both directions. When the breaker at the mill opens for any reason, it sends a pulse over the pilot wire which causes the breaker at the transmission substation to

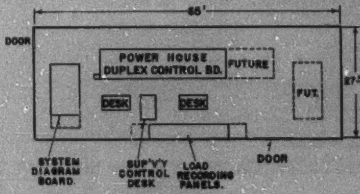


Fig. 8. Plan view of the power-dispatching control room

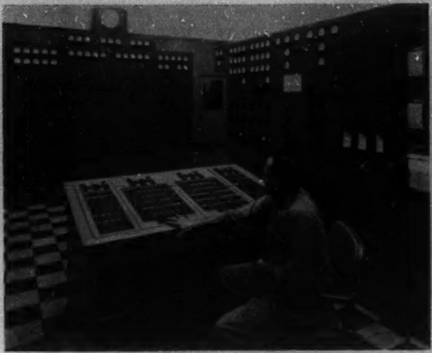


Fig. 9. View of power-dispatching control room showing supervisory control desk. System diagram board and powerhouse control board

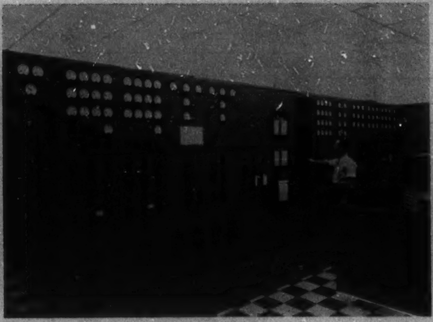


Fig. 10. Front panels of duplex powerhouse control board in power-dispatching control room

open. As long as the transmission substation breaker is open for any reason it makes it impossible to close the breaker at the mill end. The positions of all transmission substation breakers are indicated to the system operator in the powerhouse by means of supervisory control. Therefore, the system operator indirectly knows whenever a circuit breaker is open at the mill end. This arrangement makes it impossible to energize a mill bus without the knowledge of the local mill operator.

Power-Dispatching Control Room

The control of the transmission and distribution system is centralized in the power-dispatching control room, which is located at the powerhouse. Fig. 8 shows a plan view of the power-dispatching control room. Fig. 9 shows the system diagram board, control desk, and part of the powerhouse duplex control board.

The system diagram board provides lamp indications of the circuit breakers and motor-operated switches of the transmission and distribution system at Biles Island, Main, Hot Strip Mill, and Pump House Substations by means of supervisory control equipment. It also provides lamp indications for the 69-kv and 13.8-kv breakers at the powerhouse by direct connection to breaker auxiliary switch contacts. Telemetering receivers are located at the top of the system diagram board.

The powerhouse duplex control board mounts all of the control and protective relay equipment for the two generators and the 69-kv and 13.8-kv circuits located at the powerhouse. All necessary instrumentation for these circuits is also on this board. Fig. 10 shows the front panels of the entire duplex type powerhouse control board. The protective relays for the powerhouse circuits are located on the rear panels of this board. The load-recording board provides recordings of load and integrations of demand on various circuits, as will be described in detail later.

Table II. List of Supervisory Control Functions

Function	Controlled Substation Biles Island	Main	Hot Strip	Pump House
Control and Supervision				
138-kv circuit breakers	3			
Tap changers, raise-and-lower	2			
Tap changers, automatic-manual	1			
132-kv motor-operated switches	2			
69-kv circuit breakers	1	4	4	4
13.8-kv circuit breakers		13	21	
2.4-kv circuit breakers				3
Supervision Only				
Control-transfer switches	1	1	1	1
Control-transfer switch, tap changers	1			
Last stage transformer cooling	2	2	2	2
Control house door	1	1		
Annunciator	1	1	1	1
Lockout, transformer differential	2	2	2	2
Lockout, bus differential		1	1	1
Spare, unassigned				1
Total, fully-equipped	19	28	35	18
Total, space and wiring	11	12	15	12
	30	40	50	30

The control desk mounts all of the supervisory control push buttons and lamps for the remote operation of Biles Island, Main, Hot Strip Mill, and Pump House Substations. The supervisory control relays are in cases mounted on the rear panels of the system diagram board. The red and green breaker and switch-position-indicating lamps on the diagram board are in parallel with red and green lamps on the control desk for the four remotely controlled substations.

Supervisory Control System

The supervisory control system at the Fairless Works provides functions in the power-dispatching control room as listed in Table II. The control push buttons and indicating lamps for the entire system are located on the control desk shown in Fig. 9.

The supervisory control equipment is of the direct selection coded-pulse type. The control for each substation is entirely independent of the control for the other substations, with separate line wires provided from the power-dispatching control room to each substation. The supervisory control equipment operates over a single pair of wires to each substation. However, two pairs of line wires in different cables are provided from the power-dispatching control room to each substation. The control of each substation may be transferred from one to the other of these pairs of line wires from the control desk. In case of failure of the operating pair, the control is automatically transferred to the other pair.

Fig. 11 shows the rear and Fig. 12 the front of the duplex-type board provided for the control of the 69-kv circuits in the Hot Strip Mill Substation. The supervisory control relays are mounted in the cases shown on the right-hand panel of Fig. 11. The supervisory control interposing relays are at the bottom of the left-hand panel of Fig. 12. The supervisory control at the other substations is similarly located on a section of the duplex board for the control of the 69- and 13.8-kv circuits.

The supervisory control operates from 48-volt batteries in each substation and in the powerhouse. The 48-volt battery voltage on each substation is telemetered over the supervisory control line wires continuously while the supervisory control equipment is at rest. These voltages are indicated by voltmeters on the system diagram board (Fig. 9). The Biles Island 48-volt battery is also the control storage battery of the substation; the other 48-volt batteries are for supervisory control only.

Although the substations are normally controlled from the power-dispatching control room, control switches and indicating lamps are provided in the substations for local control. A manually operated control transfer switch with "supervisory" and "local" positions is provided for each circuit breaker and motor-operated switch. A circuit breaker or switch may be operated by supervisory control only when its associated control transfer switch in the substation is in the supervisory position. A common indication is provided in the power-dispatching control room for all of the control transfer switches in a substation. By means of this indication, the operator knows whether or not he has control of all of the circuit breakers and switches in a substation.

Closing operations can be performed only with a local control switch when the associated control-transfer switch is in the local position. Tripping operations can be performed by the local control switches regardless of the position of the associated control-transfer switches.

The local circuit breaker and switch-position-indicating lamps at the substation, as well as the supervisory control circuit breaker and switch-position-indicating lamps at the power station, remain operative regardless of the position of any control-transfer switches.

An *a* and a *b* auxiliary switch are used for supervisory control position indication on each circuit breaker and motor-operated switch. The *a* switch is connected to positive battery, the *b* switch to negative battery, and the supervision wire is connected to both the *a* and the *b* switches. If the *a* auxiliary switch is closed, a red light results at the powerhouse. If the *b* auxiliary switch is closed, a green light results at the powerhouse. If neither auxiliary-switch contact is closed or if the supervision wire is broken, both the red and the green lights at the powerhouse are extinguished.

Synchronizing

This system is unique in that there are not potential transformers or potential devices on the line side of any of the 69-kv transmission line breakers. All potential transformers and potential devices on the transmission system are shown in Fig. 3.

The 69-kv line potential transformers at Biles Island Substation and the 13.8-kv potential transformers on the two busses at the powerhouse must be used for synchronizing all 69-kv breakers. These potentials are made available at all locations by the use of conductors in the lead-covered telephone-type cables which provide circuits for supervisory control, pilot-wire relaying, etc. Stepup transformers are employed to provide adequate voltage for synchronizing and voltmeter indications at the various locations despite the considerable voltage drop over the several miles of telephone line wires involved. Variable resistors are connected in the various potential circuits to the synchroscope and voltmeters at each location to provide for accurate voltage adjustment. At the powerhouse, where there are two synchroscopes and two sets of voltmeters, compensating resistors are provided so that accurate indications are obtained when either or both sets of instruments are connected to remote potentials.

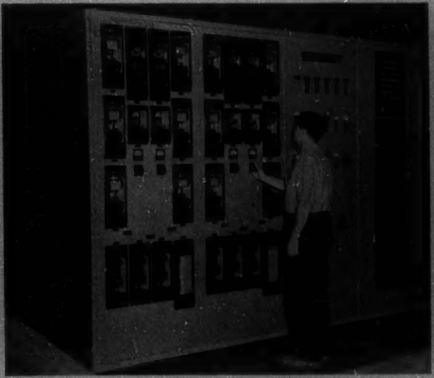


Fig. 11. View of rear panels of Hot Strip Mill Substation duplex control board

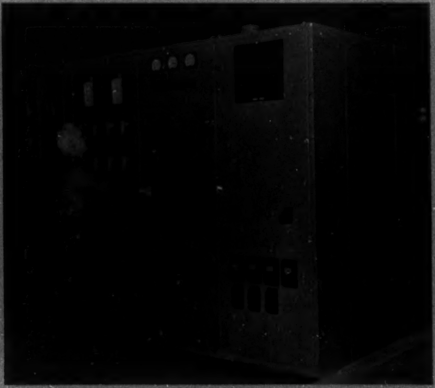


Fig. 12. View of front panels of Hot Strip Mill Substation duplex control board

The powerhouse duplex control board (at right in Fig. 9) includes synchroscope, frequency meter, voltmeters, and switches for synchronizing the circuit breakers on the generator, bus tie, and main transformer circuits. Synchronizing is conventional except for the fact that the "running" voltage for synchronizing a main transformer circuit breaker is brought in from a 69-kv potential transformer at Biles Island.

Synchronizing across any of the 132-kv or 69-kv circuit breakers of the entire system is normally performed remotely by the operator in the power-dispatching control room by means of the supervisory control system. When necessary, synchronizing also may be performed locally in the substations.

Before any 132-kv or 69-kv circuit breaker can be closed by supervisory control, the operator must first operate the selection push button for the particular circuit breaker and then operate the synchronizing push button for the particular substation. These operations set up a permissive circuit for a closing operation and also connect the proper running and incoming voltages to the synchroscope and voltmeters on the system diagram board. If synchronism conditions are then satisfactory, the operator may immediately close the breaker by means of the supervisory control equipment. If the synchronism conditions are not correct, a synchronizing transfer push button on the front of the control desk must be operated. This operation transfers all necessary circuits to the powerhouse duplex control board so that the actual synchronizing and breaker-closing functions can be performed from that location. The synchroscope and voltmeters on the duplex control board are connected in parallel with those on the system diagram board and the breaker can be closed after the powerhouse frequency has been properly adjusted by raise-lower switches on the duplex control board.

Telephone-type cable conductors for synchronizing are assigned so that voltages associated with each of the two transmission lines are carried in separate cables. In addition to the conductors for normal use, a pair of spare conductors in the other cable is provided for each voltage for use in case of failure of the normal conductors.

Tap-Changer Control

The two 30,000-kw transformers at Biles Island Substation are provided with tap-changing-under-load equipment. Each tap changer may be operated independently, or it may be operated in parallel with the other tap changer. A control transfer switch with a lock is located on the duplex control board at Biles Island Substation. When this switch is in the supervisory position, automatic or manual operation may be selected by the operator in the powerhouse by means of a supervisory control function. If manual control is selected, raise-lower control functions can be performed on each tap changer by means of the supervisory control. The position of the control-transfer switch is indicated continuously on the desk in the power-dispatching control room.

When the control-transfer switch is in the local position, automatic or manual operation may be selected for each tap changer, by means of an automatic-manual selector switch on the tap-changer control panel at the transformer. For manual raise-lower operations, raise-lower control switches on the tap changer control panel and on the duplex control board are connected in parallel for each of the two transformers.

When the automatic selection has been made either by supervisory control or by local selector switches, the raise-lower operations are controlled by the voltage-regulator relays and line-drop compensators of the two transformers.

Selection of parallel operation or independent operation of the tap changers can be made only at the tap changers, by means of a jumper and a parallel-independent selector switch provided on each tap-changer control panel. When on parallel operation, the raise-lower operations of the master tap changer cause operations of the other tap changer, regardless of the local-supervisory and manual-automatic selections in effect at the time. Out-of-step switches require that the two tap changers be on the same step before circuits can be completed for parallel operation.

Telemetering

The indicating instruments of the transmission system are mounted on the system diagram board, above the diagrams of the associated stations. Kilovolts on the 132-kv and 69-kv sides, tap-changer position, megawatts, and power factor are indicated for each transformer in Biles Island Substation. Megawatts through each transformer and kilovolts on the low-side bus sections are indicated for each of the other transmission substations. Kilovolt and megawatt indications for the powerhouse and adjacent 69-kv power House Substation are provided through local direct wiring.

A-c volts are transmitted through insulating transformers directly over the cable pairs and are indicated on suitably calibrated a-c voltmeters for the kilovolt indications. Tap-changer position indications are obtained with potentiometer-type telemetering. Rate-of-pulse type of telemetering equipment is used for the indications of megawatts. Current balance type of metering is used for power factor indications.

The load-recording board, shown in Fig. 8, provides a strip-chart megawatt recorder, and an integrating demand meter for each of the following loads:

Purchased power received through Biles Island Substation.

Power House total output to 69-kv lines.

Power House total output to 13.8-kv substations.

Hot Strip Mill Substation, total to 13.8-kv bus.

Main Substation, total to 13.8-kv bus.

Each of the five 13.8-kv substations supplied from the Hot-Strip Mill Substation.

Each of the four 13.8-kv substations supplied from Main Substation.

Each strip-chart recorder is provided with demand-meter contacts, from which an integrating demand meter is operated. The demand meter has two pointers, an "ideal" pointer that is driven at a uniform rate depending on the desired maximum demand, and an "actual" pointer which indicates the actual demand. If the actual-demand pointer passes the ideal-demand pointer, an alarm is sounded.

Each recorder (with one exception) receives the combined d-c millivolt output of two thermal converters connected in series. At the points of measurement, a thermal converter is connected to each of the two lines. Each thermal converter gives a d-c output proportional to the watts measured. For the loads measured in the power station, local wiring connects the thermal converters to the recorders; for the loads at other points, the circuits from the thermal converters to the recorders are routed through wires in the telephone-type cables.

Since the load of the Hot Strip Mill is supplied directly from the 13.8-kv bus of the Hot Strip Mill Substation, the Hot Strip Mill recorder is connected to the thermal converter d-c output circuits so that it records the Hot Strip Substation total load minus the sum of the loads of the four other 13.8-kv substations which are supplied from the Hot Strip Mill Substation.

Annunciators and Alarms

In each substation every abnormal condition not reported individually to the power dispatcher over the supervisory control system is registered on an annunciator in the substation. Each such annunciator operation causes an audible signal and a flashing green light in the power-dispatching control room. Thus the power dispatcher is kept informed of all abnormal conditions on the electric power system.

A supervision function is provided for annunciator operation in each of the supervisory controlled substations as listed in Table II. Whenever any drop operates on the annunciator in one of the four substations, a supervisory control signal is initiated by independent contacts on the annunciator bell relay. A flashing green light appears on the supervisory control desk on the escutcheon assigned to the annunciator of the substation affected. The flashing may be stopped by operation of the proper reset push button on the supervisory control desk, but the light remains green until the bell relay at the annunciator in the substation has been reset by hand.

Manual reset of the bell relay in the substation silences the annunciator bell and prepares the circuits so that subsequent operation of any drop will initiate the supervisory control signal besides ringing the bell again in the substation. Operation of an annunciator drop, whether it has been reset or not, does not prevent the operation of any other drop. Annunciator drops are reset by hand individually, but a drop will not stay reset unless the condition that caused its operation has been restored to normal.

A row of red and green lamps on the system diagram board supervises the annunciators in the 13.8-kv substations through direct-wire circuits independent of the supervisory control system. The independent contacts of each annunciator are connected to the system diagram board through a separate pair of conductors in the telephone-type cables.

The operation and supervision of these annunciators are similar to the operation and supervision of the transmission substation annunciators. The red lamps on the system diagram board indicate normal conditions. When an annunciator in one of these substations operates and closes, the independent contacts, the green lamp for that annunciator starts flashing and an alarm bell sounds. By operating a reset push button on the vertical apron of the supervisory control desk, the bell may be silenced and the green light changed from flashing to steady green, without preventing the reception of later signals in case another substation annunciator should operate. The steady green light continues until the substation annunciator has been reset manually, at which time the light on the system diagram board changes back to red. One set of the red and green lamps on the system diagram board supervises an additional annunciator in the Pump House Substation, through direct-wire connections just described. This annunciator is operated by the tripping of any of the 2.4-kv breakers not controlled by supervisory control.

Generally, annunciator drops are provided for such abnormal conditions as require local investigation and action in the substations. For each 69-kv transformer (except at the pump house) drops are provided for high or low pressure in gas spaces, high temperature of oil or hot spot, low level of oil, operation of pressure relief device, failure of oil pump, and loss of auxiliary power for pumps and fans. The pump house 69-kv transformers have alarm contacts only for hot oil and for loss of control voltage. Drops are provided for loss of normal supply of substation auxiliary power, for ground or low voltage on substation battery circuits, and for failure of battery charger. Each 132-kv and 69-kv circuit breaker is supervised by a drop for low air pressure in the pneumatic operating mechanism. Supervision relays on the pilot wires for protective relaying are connected to operate annunciator drops in the event of pilot-wire failure.

The power-dispatching control room receives indication, either by supervisory control or on the power house control board, whenever any circuit breaker on a feeder to a 13.8-kv substation trips. Tripping of the other circuit breakers in these substations is supervised by individual drops on the substation annunciators. Drops also are provided for such other abnormal conditions as apply in these substations.

Cables for Pilot Wires, Supervisory Control, Etc.

Two lead-covered, telephone-type cables, with 18-gauge conductors run in the underground telephone duct system, connect the powerhouse with all 69-kv substations by a routing which goes from the powerhouse to the central telephone exchange and then divides. One group of cables goes to Biles Island Substation, and the other group goes to the Main Substation, then to the Hot Strip Mill Substation, and then to the Pump House Substation. The cables provide conductor pairs for the following electric power system functions:

1. Pilot wires for pilot-wire relaying.
2. Supervisory control line wires.
3. Telemetering.
4. Synchronizing voltages.
5. Low-frequency tripping.
6. Thermal converter wires to load recorders.
7. Remote indication of annunciator operation.

The two cables are distinguished by *A* and *B* designations. Functions associated with the *A* transmission circuits, transformers and busses are assigned to pairs in the *A* cable, those associated with *B* circuits are assigned to pairs in the *B* cable. Supervisory control line wires are provided in both cables, with provisions for automatic and manual transfer. In addition to the pairs that are connected and used for synchronizing voltages and low-frequency tripping, a spare pair in the other cable is reserved for each of these circuits. These spare pairs are connected to small double-throw switches at the switchboard terminals, so that manual transfer may be made readily and correctly, in the event of failure of the pairs that are normally used.

Pilot-Wire Protection

Pilot-wire protection is required because the high ground resistivity and relatively high ground fault currents cause large differences in station ground potentials during ground faults. Because of the routing of the cables, there are pilot wire pairs in the same cable which terminate at different stations, so that all pairs are not subjected to the same difference in ground potential. A special study indicated that all pairs subjected to these conditions should be equipped with gas-filled protector tubes to bring all pairs in a cable to station ground potential if the potential exceeds 300 volts. Pairs used for pilot-wire relaying have drainage reactors in addition to protector tubes to prevent the latter from short-circuiting out the a-c signal during faults. On the pilot-wire relay circuits protecting the 13.8-kv feeders from the powerhouse, 2-winding neutralizing transformers were used because all wires in the same cable terminated at the same substation. The pilot wire circuits for the other 13.8-kv feeders did not require protection because the ground mats at the mill substations are tied together, eliminating differences in substation ground potential.

Interlocks

Service and equipment are safeguarded by electrical interlock to prevent closing circuit breakers out of synchronism, overloading 13.8-kv feeders by load transfer, and opening or closing air switches under load.

Normally the system operates with the *A* and *B* circuits tied together through 69-kv tie circuit breakers in the Biles Island and Power House Substations. At the other 69-kv substations, the *A* and *B* circuits are normally tied together through the *A* and *B* transformers, the transformer circuit breakers and the 13.8-or 2.4-kv bus-tie circuit breaker. To prevent completing a tie through the transformers when out of synchronism, each of these stations is provided with a synchroverifier energized from potential transformers connected to the low-side terminals of the respective *A* and *B* main transformers and with voltage relays energized from potential transformers connected to the *A* and *B* sections of the low-side bus. A transformer circuit breaker cannot be closed unless either the synchroverifier shows synchronism between the two main transformers or the bus section connected to the circuit breaker is de-energized. The bus-tie circuit breaker cannot be closed unless either the synchroverifier shows synchronism, or one of the bus sections is de-energized and the adjacent transformer breaker is open.

In the Main and Hot Strip Mill Substations, wherever two 13.8-kv feeder circuit breakers on opposite sides of the bus tie circuit breaker supply a pair of feeders to the same distribution substation, interlocks are provided to prevent closing either one of these feeder circuit breakers unless the other feeder circuit breaker of the pair is open, or unless the bus-tie circuit breaker is closed. Closing or opening of the bus-tie circuit breaker is prevented unless one circuit breaker, of each pair of feeders supplied from the bus, is open. Thus the 13.8-kv feeders are protected against being overloaded by interchange current in case of abnormal differences between the *A* and *B* circuits.

In the Main, Hot Strip Mill and Pump House Substations, the transformer 69-kv disconnecting switch cannot be opened or closed unless both the transformer low-side circuit breaker and the 69-kv *AB* tie-disconnecting switch are open. The 69-kv *AB* tie-disconnecting switch cannot be opened or closed unless the 69-kv disconnecting switch and the low-side circuit breaker, of one or the other main transformers, are open. Thus these disconnecting switches are prevented from opening or closing load current.

In addition to satisfying the conditions for prevention of closing out of synchronism, one of the following three additional conditions must be satisfied before a transformer low-side circuit breaker can be closed:

1. The 69-kv disconnecting switch of the transformer and the 69-kv *AB* tie-disconnecting switch must both be open. This setup permits testing with circuit breaker protection after an outage.
2. When placing the transformer in service with normal connections, the transformer 69-kv disconnecting switch must be in the closed position and the transformer differential transfer relay must be in the normal, nontransferred position.
3. When placing a transformer in service on the other 69-kv circuit, the 69-kv disconnecting switch of the transformer must be open, the 69-kv *AB* tie-disconnecting switch must be in the closed position, and the transformer differential transfer relay must be in the transferred position.

Conclusion

The a-c power system at the Fairless Works is one of the largest installations of its kind to have been designed, manufactured, and installed in so short a time. Approximately 2 years after the first network calculator study was made on proposed designs of the electric power system (Fig. 2), the construction of the various mills and the installation of all electric power equipment had been completed. This included the five large outdoor substations, the 69-kv high-tension line and the powerhouse, in addition to substations located in every mill area.

No Discussion

Induction Heating Steel with 60 Cycles

C. D. KRAMER
ASSOCIATE MEMBER AIEE

A NEW field for the application of induction heating has developed. High-frequency heating of metal is already accepted as an indispensable tool of industry for surface heating and for through-heating small pieces. The steel industry is now turning its attention to lower frequencies for heating large sections to forging temperatures, approximately 2200 degrees Fahrenheit (F). The heating of steel for forming is a large part of steel manufacture. The methods used over the years have gone through an evolution. With the development of high-speed rolling and forming techniques, a need for faster heating methods has arisen. Low-frequency induction heating seems to provide a good answer to this need.

In the low-frequency range, commercially available frequencies, usually 60 cycles, are most popular, primarily because of the lower operating and first costs. In order that full advantage be taken of 60-cycle induction heating of steel, it is necessary to know what happens during the heating cycle. It is also necessary to understand why certain sizes of steel sections are better than others for the application of 60-cycle induction heating.

Discussion

INDEX RATIO

Over-all efficiency of heating a charge depends upon radiation losses from the charge and electrical efficiency. Electrical efficiency is dependent upon the ability of a charge to absorb energy. It is very critically affected by a quantity called index ratio, where for round sections

$$\Delta=\frac{{}^1/_2 \text{ thickness of billet}}{\text{depth of penetration}} \tag{1}$$

and for slabs

$$\Delta=\frac{\sqrt{2}\,{}^1/_2 \text{ thickness of billet}}{\text{depth of penetration}} \tag{2}$$

The index ratio, then, is a primary indication of whether or not a charge of a given cross section can be heated with a given frequency.

DEPTH OF PENETRATION

For a billet of uniform permeability and resistivity, magnetic flux intensity, hence current density, decreases exponentially from the surface of the billet toward the center, as shown in Fig. 1. This is true providing the index ratio is of the order of at least 6. The depth of penetration δ is defined as the depth at which the current density is 1/e times or 36.8 per cent of its surface value. Heat generated is proportional to the square of the current. Hence, within the surface layer of depth δ $[(1.0)^2-(0.368)^2]$ or 86.5% of the heat is generated. Providing the billet thickness is large compared to the depth of penetration

$$\delta=\frac{1}{2\pi}\sqrt{\frac{\rho 10^9}{2f\mu}} \tag{3}$$

or

$$\delta=3,570\sqrt{\frac{\rho}{f\mu}} \text{ centimeters}$$

or

$$\delta=1,400\sqrt{\frac{\rho}{f\mu}} \text{ inches}$$

where

ρ = resistivity of material at temperature, ohm-centimeters (ohm-cm)
f = frequency, cycles per seconds (cps)
μ = permeability of material at temperature

Fig. 2 shows that the resistivity of low carbon steel increases steadily from about 20×10^{-6} ohm-cm at room temperature to about 120×10^{-6} ohm-cm at the Curie point. The Curie point, which occurs in low carbon steel at about 1,400 F, is the point where steel undergoes a phase change and becomes nonmagnetic. From the Curie point to the melting temperature resistivity increases very slightly.

The permeability of steel is difficult to fix numerically because it changes with flux density; it also changes with temperature. At the Curie point, the permeability may be assumed to drop to unity. For the purposes of this discussion, the permeability of low carbon steel below Curie may be assumed to be 2,000. Equation 3 may be used only when permeability is constant throughout the cross section of the billet. This occurs when the entire billet is above the Curie temperature. It also occurs when the maximum temperature in the billet is below Curie.

The resistivity, of course, also changes for different temperatures and, for accuracy, may be put into the equation only when billet temperature is uniform. This is particularly important below Curie where resistivity changes rapidly with temperature.

For low carbon steel at 70F take $\rho=20\times10^{-6}$ ohm-cm; $f=60$ cps; $\mu=2,000$; then $\delta=0.018$ inch.

For low carbon steel above Curie temperature take $\rho=120\times10^{-6}$ ohm-cm; $f=60$ cps; $\mu=1$; then $\delta=2$ inches.

The latter figure is only approximate for temperatures between Curie and forging because, as stated before, resistivity changes slightly in this range.

DESCRIPTION OF THE HEATING CYCLE

When power is first applied to a magnetic steel billet, resistivity and perme-

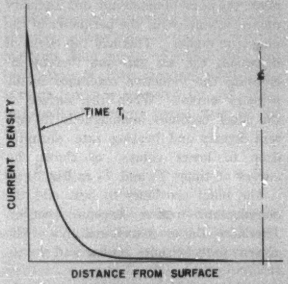


Fig. 1. Current distribution in a billet of uniform permeability

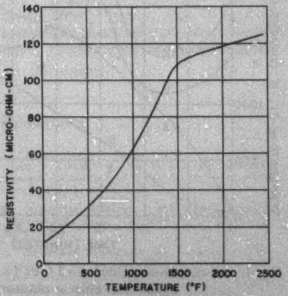


Fig. 2. Resistivity versus temperature for 0.1-per-cent carbon steel billet

---

Paper 54-465, recommended by the AIEE Electric Heating Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Middle Eastern District Meeting, Reading, Pa., October 5–7, 1954. Manuscript submitted July 7, 1954; made available for printing August 11, 1954.

C. D. KRAMER is with the West Penn Power Company, Pittsburgh, Pa.

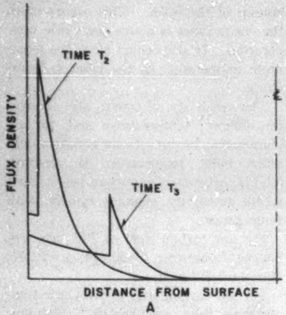


Fig. 3(A). Flux distribution in a magnetic steel billet as the nonmagnetic region moves in from the surface

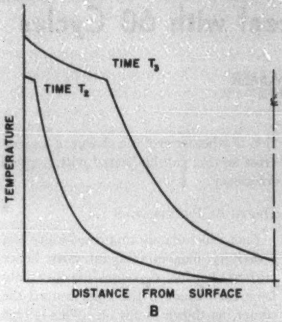


Fig. 3(B). Temperature at times T_2 and T_3 corresponding with times T_2 and T_3 of Fig. 3(A)

ability being constant throughout the billet, current density assumes the exponential distribution shown at time T_1 in Fig. 1. As the steel begins to heat, the resistance of the current path increases; see Fig. 2. The surface heats rapidly to the Curie point. As the surface becomes nonmagnetic, the point of maximum flux intensity moves inward with the perimeter of the magnetic region. This has the effect of increasing the air gap and thereby increasing the inductive reactance in the primary circuit. When the surface of the steel becomes nonmagnetic its current density and heating rate abruptly drop to lower values, as shown by curves at times T_2 and T_3 in Fig. 3.

The billet continues to heat and the nonmagnetic region becomes deeper. Therefore, the cross-sectional area of the current path becomes greater and the resistance to current flow becomes less. Current increases causing the power to increase; see Fig. 4. This phase of behavior continues until the center of the billet becomes nonmagnetic. When this happens, there is no longer a magnetic core providing good flux linkage with the billet and the power abruptly decreases. Now all portions of the billet are above the Curie temperature and the depth of the penetration formula given once again applies. This depth was calculated at 2 inches for the case of 60-cycle power.

Index Ratio and Efficiency

The index ratio, when it applies for a given billet, is determined by the frequency. Higher frequency results in higher index ratio. For a given frequency and a given temperature, billet thickness determines the index ratio. As can be seen from Fig. 5, if the index ratio of a round section becomes less than 2.5, the electrical efficiency falls off rapidly. This is also approximately true for slabs. The shapes of the curve of electrical efficiency versus index ratio for rounds and slabs differ slightly. The magnetic property of a steel billet causes the index ratio to be much higher than it is for the same billet in the nonmagnetic state. The magnetic property also causes good flux linkage with the billet. Therefore, the limiting situation occurs, if it occurs at all, when the steel billet is nonmagnetic.

The depth of penetration for low carbon steel above Curie was calculated to be 2 inches, using 60-cycle power. To insure an index ratio of 2.5, the diameter of a round billet would have to be, from equation 1

$$\begin{aligned}\text{diameter} &= 2\Delta\delta \text{ inches}\\ &= 2\,(2.5)\,2 \text{ inches}\\ &= 10 \text{ inches}\end{aligned}$$

For a slab the thickness would have to be, from equation 2

$$\begin{aligned}\text{thickness} &= \frac{2\,\Delta\delta}{\sqrt{2}}\\ &= \frac{2\,(2.5)\,2}{\sqrt{2}} \text{ inches}\\ &= 7.07 \text{ inches}\end{aligned}$$

The curves in Fig. 6, plotted from test data, show that the 8-inch and the 13-inch square billets took approximately the same amount of power when they were magnetic and their index ratios were well above 2.5. When they became nonmagnetic, the power assumed by the 8-inch piece dropped well below that assumed by the 13-inch piece. The above Curie index ratio for an 8-inch round is 2.0, which puts the efficiency below the knee of the curve of Fig. 5. Data show that equation 1, given for rounds, yields approximate index ratios for squares.

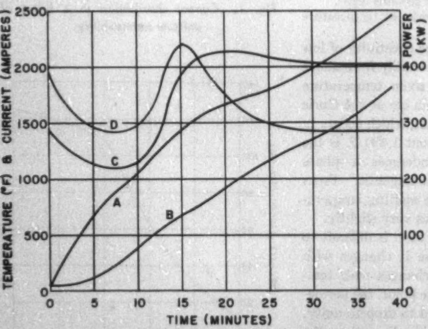


Fig. 4. Heating characteristics of a 13- by 13- by 32-inch 0.1-per-cent carbon steel billet at constant voltage

A—Surface temperature
B—Center temperature
C—Primary current
D—Power input to coil

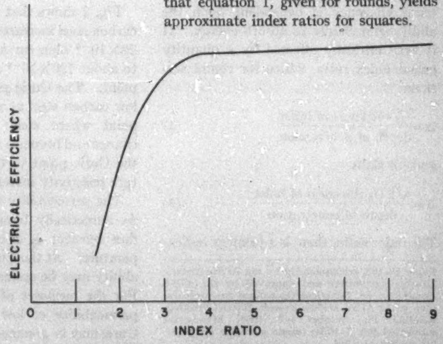


Fig. 5. Electrical efficiency versus index ratio for a solid cylindrical billet

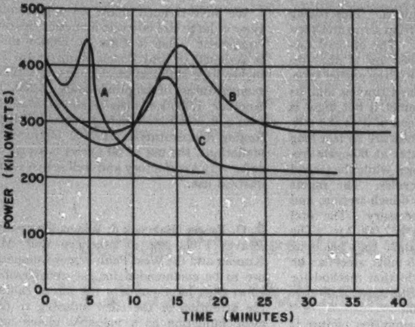


Fig. 6. Comparison of power curves for three cross sections of 0.1-per-cent carbon steel

A—8 by 8 inches
B—13 by 13 inches
C—6 by 18 inches

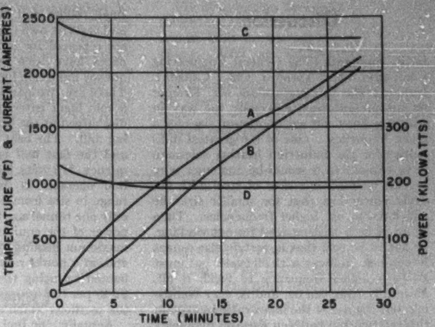


Fig. 7. Heating characteristics of a 7- by 8- by 32-inch nonmagnetic stainless steel billet

A—Surface temperature
B—Center temperature
C—Primary current
D—Power input to coil

From equation 2, the index ratio for a 6- by 18-inch slab is 2.1; see Fig. 6. As in the case of the 8- by 8-inch piece, whose index ratio is 2, the power taken by the slab drops to its lowest value when the billet becomes nonmagnetic. However, this drop is not nearly so great for the slab as for the 8- by 8-inch piece. This illustrates the fact that the knee of the electrical efficiency curve for slabs occurs at a lower value of index ratio than it does in the case of squares or rounds. It is possible to heat a steel billet as small as 3 inches in diameter to forging temperature with 60-cycle power although efficiency would be quite low. For pieces smaller than 3 inches, higher frequencies must be used above the Curie point.

NONMAGNETIC STEEL

A comparison of the curves of Figs. 4 and 7 for magnetic and nonmagnetic steel shows that kilowatt- and ampere-curves for nonmagnetic steel do not peak as they do for magnetic steel. This is, of course, because the permeability does not change appreciably as the billet heats.

The temperature difference between the surface and the center is less for the nonmagnetic steel. This is because the billet takes a little longer to heat. Also, the heat is being generated deeper within the billet during the initial phases of the heating cycle. By watching the tops of the billets while they are heating, this difference in heating depth may be easily observed. A magnetic piece shows a well-defined line between red and black. A nonmagnetic piece exhibits a fairly gradual color change from surface to center.

MAXIMUM SIZE

When the index ratio is very large, the heating is being done very near the surface of the billet. The interior portion of the billet receives heat chiefly by thermal conduction. This is a relatively slow process. Time required for heating the center to temperature is great. The result is a reduction in over-all heating efficiency because of increased radiation losses from the surface of the billet. Billets and ingots normally produced in steel mills are not considered very large in this sense. They do not usually exceed 36 inches in thickness, with index ratios from 9 to 13. This is within the range of good heating efficiency for 60 cycles.

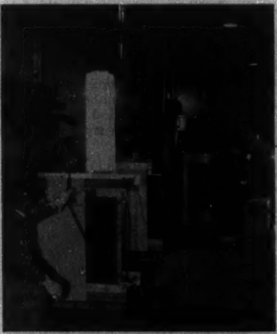


Fig. 8. A hot 8- by 8- by 32-inch carbon steel billet being removed from a test coil

HYSTERESIS

In a laminated transformer core, hysteresis power is large enough with respect to eddy current power to be a factor worthy of consideration. In the case of heating large solid masses of metal, with the high values of power normally used, eddy current power is so much greater that, ordinarily, hysteresis power may be ignored.

Conclusions

Round or square steel sections larger than 10 inches across, and slabs larger than 7 inches thick may be heated with 60-cycle power with very good electrical efficiencies of the order of 70 or 80 per cent. Heating round or square steel sections between 3 and 10 inches across, and slabs between 3 and 7 inches thick may be economically justifiable. As pieces become smaller than 10 and 7 inches, efficiencies become lower. Smaller pieces require higher frequencies above the Curie point.

References

1. HEAT TREATMENT OF STEEL BY HIGH FREQUENCY CURRENTS, G. Babat, M. Losinsky. *Journal*, Institution of Electrical Engineers, London, England, vol. 86, 1940, pp. 161–168.

2. INDUCTION HEATING (book), N. R. Stansel. McGraw-Hill Book Company, Inc., New York, N. Y., 1949.

3. STANDARD HANDBOOK FOR ELECTRICAL ENGINEERS (book), A. E. Knowlton. McGraw-Hill Book Company, Inc., New York, N. Y. 1941, pp. 1,774–79.

4. INDUSTRIAL ELECTRIC FURNACES (book), V. Paschkis. Interscience Publishers, Inc., New York, N. Y., vol. 2, 1948, pp. 209–32.

Discussion

O. R. Summers (The General Engineering Company, Ltd., Toronto, Ont., Canada): I should like to congratulate Mr. Kramer on his presentation, especially the way in which he has dealt with optimum sizes for a given frequency. One of the greatest difficulties for the induction heating salesman is to explain to a would-be customer why he can heat certain sizes of steel with 60-cycle power but that for smaller sizes he will have to use higher frequencies. Then again, if he is heating steel for hot working, he has to be told that in certain size ranges he can heat halfway with 60 cycles but must go to a higher frequency to finish it off. Naturally, this can be, and is, very confusing, but the use of the index ratio to explain this relationship between size and frequency is very helpful.

It has been pointed out that the index ratio for efficient heating should be a minimum of 2.5. How much it is above this figure is not so important within limits but, in practice, it is better when selecting a frequency for through heating to obtain an index ratio as low as possible or, in other words, to select the lowest standard frequency that will be electrically efficient. This frequency will give the deepest penetration and as we are aiming for through heating, the greater the penetration the greater the heating rate and the lower the thermal gradient. For surface hardening, where the skin effect is desired, these arguments do not hold, but that type of application is not being considered in this paper. During the heating cycle, the index ratio is changing all the time as the resistivity will change with temperature and, in the case of magnetic steel, the permeability changes very severe through the Curie point. This has led some manufacturers to develop frequency units which use 60 cycles for half the heating and a higher frequency for the remainder. Thus a 3-inch-square steel billet can be very efficiently heated to the Curie point with 60 cycles but after this the application becomes very inefficient.

Thus, a heat tunnel is constructed in two parts, the first one operating at 60 cycles which raises the billet temperature to 1,400 F. The billets pass continuously on into a 600-cycle tunnel which raises them to forging temperature. By this means, the index ratio is stabilized for the full heat cycle in spite of the great change in permeability.

The power factor is poor for steel. For the magnetic condition, it may be around 50 or 60 per cent, but for steel above the Curie point, it may go down to 20 or 25 per cent. During the heat cycle of magnetic steel, the power factor will be changing all the time and at Curie point it will change very severely. If one billet is being heated in one coil, this means that capacitors would have to be switched into circuit to take care of the decreasing power factor. Of course, this does not have to be done as it makes no difference to the heating effect of the coil but naturally makes the line current very high and the power authorities object strongly. We overcome this difficulty by using a long tunnel where billets move through continuously so that each coil of the tunnel is looking at a constant load all the time.

We have recently put into commission a dual-frequency induction billet heater for a small rolling mill in Canada. This rolling mill casts its own ingots from scrap and then are reheated for rolling. The oil-fired furnaces have been replaced by a dual-frequency induction furnace of the pusher type. There are two parallel heat tunnels and, as a cold ingot enters a tunnel, a hot ingot is discharged at the other end of it to the rolling mill. The heat tunnels are 35 feet long and the first half operates at 60-cycle frequency, 575 volts, 3-phases, while the second half operates at 540 cycles. The ingots range in size from 4- to 6-inch-section, and only one tunnel size is necessary. The total power of the equipment is 7,000 kw. The maximum heating rate is 25 tons per hour and, at a power rate of 5 mills, steel can be heated by using the induction method for $1.43 per ton as opposed to some $3.00-per-ton methods which use oil.

Actually, the fuel consumption picture is very variable because of local conditions and, where very cheap supplies of gas are available, the induction method in fuel alone is more expensive. However, the story with induction rests much more on the large savings due to lack of scale. In the installation mentioned, the ingot takes approximately 20 minutes to heat by induction. When rolling a 1-inch-diameter bar, 6-per-cent increase yield was experienced over previous heating with oil-fired furnaces.

Mr. Kramer worked out some examples of sizes and index ratio in the case of steel below and above the Curie point. For steel above the Curie point, for a frequency of 60 cycles, a depth of penetration of 2 inches is obtained. We have an induction billet heater operating in Montreal inches on the heating of 155-millimeter shell billets for forging. These billets are 6 inches square and the equipment is all 60 cycles right up to forging temperature.

The knee in the index ratio efficiency curve does seem to be a little erratic, which means that in the end each case must be tested separately, particularly fringe cases. This is particularly true of hollow cylinders. The wall thickness would appear to be the critical dimension but, in practice, the diameter plays a considerable part.

In conclusion, I would like to add that, although Mr. Kramer has used steel to indicate the points he makes, the same laws apply equally to all the nonferrous metals, Al Cu, Cu Ni, brass, etc., and, in this field, because of low resistivity, 60 cycles can be used down to diameters of 2 to 3 inches with very great success. In the case of nonferrous materials we have no permeability problem and 60-cycle heating for aluminum extrusion is widely established. A set of figures giving the theoretical minimum diameters for optimum efficiency with 60-cycle frequency are as follows:

Steel below Curie, $2^1/_4$ inches
Steel above Curie, $10^1/_2$ inches
a-c copper, $2^7/_8$ inches
Brass, $3^3/_4$ inches

B. E. McArthur (Magnethermic Corporation, Youngstown, Ohio): Mr. Kramer is to be congratulated on his fine paper. It is a pleasure to see others presenting papers and developing markets in the 60-cycle induction heating field. West Penn Power has shown remarkable foresight and is to be congratulated on their work in the field.

We have recently made several installations of large size 60-cycle induction heating equipment. One of these, a 400-kw unit, is designed to heat 7-inch diameter by 28-inches long stainless steel billets to forging temperatures of approximately 2,250 F. Another 1,000-kw installation is used to heat 5, 6, and 8-inch-diameter steel billets to forging temperatures. In both of these installations, the use of 60 cycles has proved to be very satisfactory and well suited to industrial use.

S. O. Evans (Babcock & Wilcox Company, Beaver Falls, Pa.): I believe that Mr. Kramer and the West Penn Power Company are to be commended for the considerable effort they have put into the job of calling the attention of the steel industry to 60-cycle induction as a valuable bloom and billet heating tool.

My comments will be concerned with the operating significance of this method of heating in the extrusion department of my company. To clarify my comments, a brief description of the process is in order.

Billets of 7 or $8^1/_4$ inches in diameter and 14 to 24 inches long are heated in a gas-fired furnace to just under 2,300 F using a protective coating to prevent scale. They are then pierced in a vertical 500-ton press, providing a hollow billet for the extrusion of tubes. In the course of conveying, piercing, and inspection $1^1/_2$ to 3 minutes may elapse, part of which time the billet is in contact with relatively cold tools. It loses several hundred degrees of temperature, particularly near the outside surface. We use 60-cycle induction to replace this lost heat.

The significant aspects of induction heating relate to the quality of the finished product. First, extrusion of tubing requires uniform heating around the axis because a cold and a hot side would cause a discrepancy in the flow of the steel resulting in eccentricity of the tube. This uniformity produced by induction is highly satisfactory. Second, the heating method should either protect the billet from air or be fast enough so that die-destroying scale does not get a chance to form. Induction heating adapts very well to the provision of inert atmosphere. Our equipment is capable of being run with such an atmosphere, but we have found that it is unnecessary to provide it on the major portion of our work because the small reheating time, about $1^1/_2$ minutes does not cause a significant scale on the stainless steel being processed. Third, the means of heating should be such, in terms of time and corrosive conditions, that the extruded surfaces will be good. We have found that the induction reheat has given us superior outside and inside surfaces. Fortunately, the inside surface which would be most difficult to repair by grinding is particularly good.

In general, after more than 6 months of production operation we have found the equipment highly satisfactory. There are still minor refinements to be done in our operation, but the fundamental operation is good.

C. D. Kramer: It is interesting to note from the discussions of Mr. Summers and Mr. McArthur that 60-cycle induction heating has progressed far beyond the theoretical stage. Installations are going in and operating experience is being gained.

The inherent advantages mentioned by Mr. Evans have a real dollar value which must be added in when comparing operating costs of induction and other methods.

One additional thing which should be mentioned is the effect of saturation of the steel. With high values of flux density, the depth of penetration will be greater than calculated. This occurs when the flux-density peak of Fig. 3(A) would be above the saturated value for the steel. This limits the minimum size of magnetic steel which can be heated to the Curie point to about $1^{1}/_{4}$ inches for values of power normally used with 60 cycles. For larger sizes, this is apparently a negligible factor.

Use of induction heating is of course not limited to billet shapes. Coil configurations can be built for heating a great variety of pieces. With an understanding of the fundamentals given in the paper, the user can undoubtedly see many applications in his plant. Equipment manufacturers will co-operate with user in developing processes around induction coils.

Carbon-Pile Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation

H. H. C. RICHARDS
ASSOCIATE MEMBER AIEE

FOR many years during the initial use of the carbon-pile regulator for control of aircraft generator systems, the adjustment of the regulator was considered to be extremely difficult. This is no longer true. Because of the instability of early regulator designs, there was no attempt to adjust the regulator at its theoretically optimum adjustment point. This was brought out in previous papers[1-3] on the carbon-pile regulator which cover its construction, component characteristics, principles of operation, and adjustment. References to adjustment procedures have generally pointed out that the system would become unstable over a large portion of the adjustment range and that this region must be avoided. The result was that the instructions given for locating the stack screw adjustment point were by trial and error in the narrow region of stability. Fig. 1 shows the relation between the stable region of early regulator designs and the optimum adjustment point. At the time this was the only method but, in the design in Fig. 2, the instability during adjustment no longer exists because of the addition of a mechanical damping feature[4] and an electric stabilizing circuit; see Fig. 3(A). With the stability problem solved, the adjustment procedure may be reconsidered. It is the purpose of this paper to show that the original adjustment procedure by the cut and try method can be replaced by a more direct method. It will be shown that the theoretically optimum adjustment point falls at a specific point on the adjustment curve. In accomplishing this, it will be pointed out that the electromagnet core position, also previously located by a trial and error method, may be correctly located during assembly when an armature calibration measurement is made.

The component characteristics and the general theory of operation will be briefly restated since the calibration and adjustment methods are developed from them. For simplicity, the effect of hysteresis and friction are not included in the curves. Their effect on the results are shown in the Appendix. The figures in the Appendix are from test data on the regulator and serve to show the correlation between the theory developed in the text and the actual results. In addition, the effects of a few environmentals on the regulator are given in the Appendix.

Regulator Construction

An external view of a regulator is shown in Fig. 2. This particular regulator is for controlling the voltage of a 30-volt d-c generator. The regulator is suitable for use with an a-c generator by a slight change in the internal construction and by the addition of a transformer and rectifier. The discussion in this article will deal with the d-c application only.

A cutaway view of this same regulator is shown in Fig. 4. Starting on the left, the first basic component is an electromagnet circuit made up of a magnetic yoke and a coil. The second basic component is an armature assembly made up of a star shaped cantilever spring attached to an armature plate which is actually a section of the electromagnet. For simplicity, however, the armature plate will be considered as a part of the armature assembly. The armature is magnetically attracted by the electromagnet. This action is opposed by the spring. The third basic component is the carbon pile or, as it is sometimes called, the pile or stack. The stack is made up of a number of carbon disks which resemble washers. The stack is brought into position by the large screw found on the extreme right and is acted upon by the difference in force between the magnet and spring. The spring, operating at a higher force than the magnet, keeps the stack under pressure of varying degrees depending upon the needs of the generator. The needs of the generator are sensed by the coil in the magnet. The fourth basic component is the base assembly. The base acts as a support for the foregoing three components and for resistors, rheostats, terminals, and wiring. The wiring scheme for a simple d-c regulator is shown in Fig. 3(A) and the wiring for the regulator-generator system is shown in Fig. 3(B).

Component Characteristics

The electromagnet for the regulator is provided with an adjustable core. The core is used to adjust the pulling strength of the electromagnet for a given value of coil current. This is necessary to offset the normal manufacturing variations in the electromagnet and to calibrate it with respect to the armature spring. With the coil turns and coil voltage held constant, the effect of varying the core position on the magnet force at various air gaps is shown in Fig. 5. The magnet face is the zero reference for the air gap between the magnet face and armature plate and for all other measurements except the change-in-length measurements of Fig. 6. For a given position of the core and with the coil turns and circuit resistance fixed, the effect on the magnet

Paper 54-356, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted June 15, 1954; made available for printing August 5, 1954.

H. H. C. Richards is with the Westinghouse Electric Corporation, Lima, Ohio.

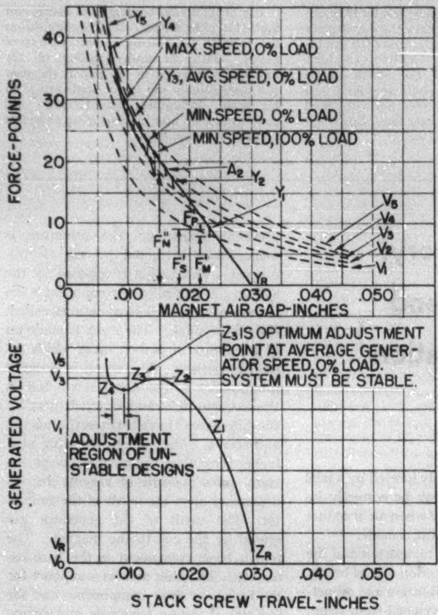


Fig. 1 (left). Determination of adjustment curve and the optimum stack-screw adjustment point

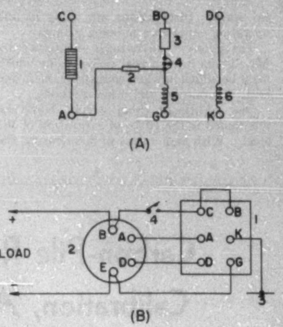


Fig. 3 (right). (A) Wiring diagram of regulator showing 1. carbon stack, 2. stabilizing resistor, 3. swamping resistor, 4. rheostat, 5. operating coil, and 6. equalizer coil. (B) Generator-regulator system wiring showing 1. regulator, 2. generator, 3. equalizer bus, and 4. field switch

force at various air gaps and applied voltages is shown in Fig. 7. In actual operation under a fixed set of environmental conditions, the core position, coil turns, and circuit resistance are fixed; therefore, the magnet characteristics of Fig. 7 are the ones to be used.

The characteristics of the armature assembly, being a spring assembly, can be represented by the linear force-deflection characteristic of a cantilever beam over the first part of its deflection, as shown from right to left in Fig. 4. Over the last of the spring deflection the armature is designed such that the length of the beam is shortened by the abutment ring on which it rests. The resulting shortening of the beam increases its stiffness, producing the curve shown in Fig. 8. Since the deflection of the armature assembly is toward the magnet face, the curve is plotted accordingly as a convenience in plotting the spring curve in Fig. 9. The normal plotting of the spring curve, when it is considered as a separate item, is shown in the small insert in Fig. 8.

The characteristics of the carbon stack involve two quantities which vary with the force exerted axially along the stack. As the force is varied, the stack will change its length and electrical resistance simultaneously. Both of these characteristics are nonlinear and are most conveniently represented on semilog and log-log graphs respectively. However, for this discussion, rectangular co-ordinates must be used and both curves must be represented in one graph. Fig. 6 is a graph of the force-length-resistance characteristic on rectangular co-ordinates.

The three basic component charac-

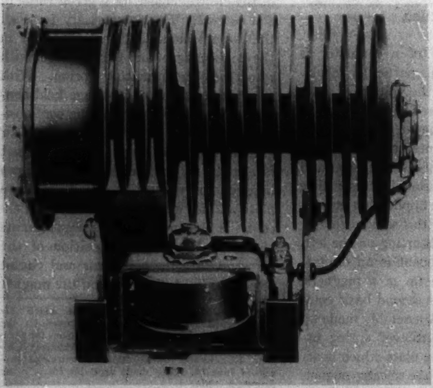


Fig. 2. External view of a carbon-pile regulator

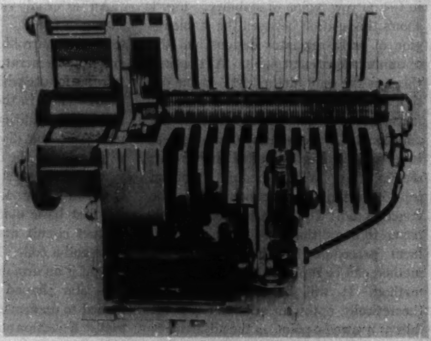


Fig. 4. Cutaway view of a carbon-pile regulator

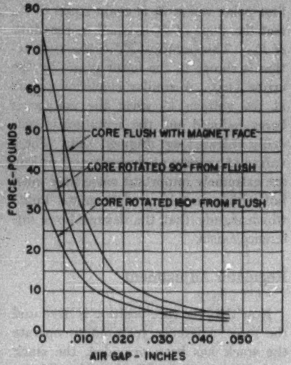


Fig. 5. Electromagnet force air gap curves for various core positions. Coil turns and applied voltage are held constant

teristics are not completely reversible because of their respective hysteresis loops. To avoid unnecessary complications, the hysteresis loops will not be considered since they are not needed to present the basic theory of operation; however, they do present limitations in the accuracy of operation of the actual device.

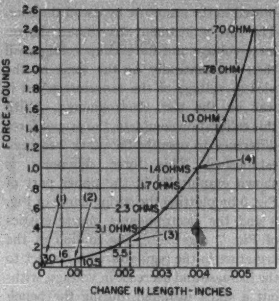


Fig. 6. Carbon stack force-resistance-length curve showing the stack requirement at 1. maximum speed, 0-per-cent load, force = F_p'', 2. average speed, 0-per-cent load, N ohms, force = F_p', 3. minimum speed, 0-per-cent load, and 4. minimum speed, 100-per-cent load

General Regulator Theory

To maintain a constant terminal voltage of a generator over its speed and load range, a controlled variable resistance must be placed in series with the generator field. The variable resistance in this case is the carbon stack which is acted upon by the spring in the armature assembly and the electromagnet. The spring exerts a force to compress the stack and the magnet exerts a force according to the signal from the generator to reduce the pressure on the stack. The armature spring force

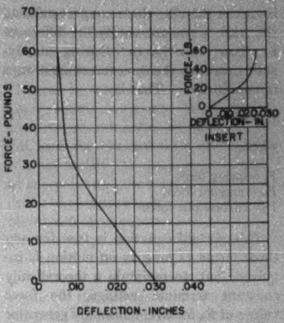


Fig. 8. Spring force deflection curve. Insert shows normal plotting of curve with respect to the intersection of the axes instead of the electromagnet face

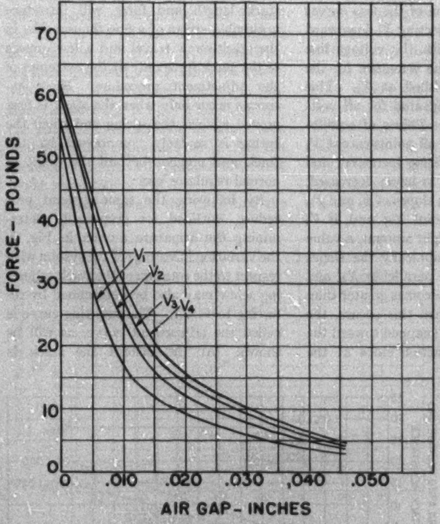


Fig. 7. Electromagnet force air gap curves for various applied voltages. Coil turns, circuit resistance and core position are held constant. V = applied voltage

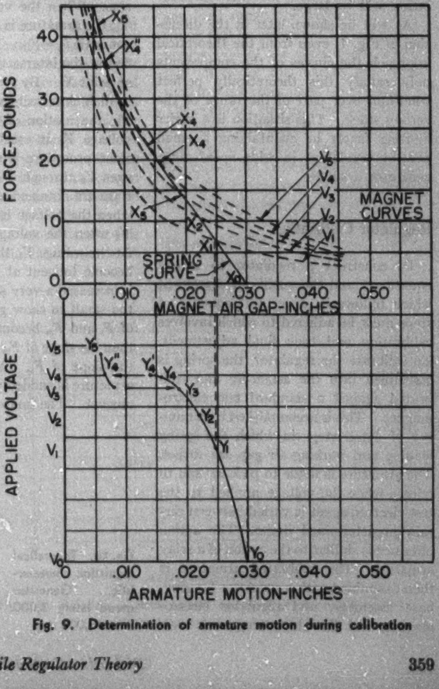


Fig. 9. Determination of armature motion during calibration

is greater than the electromagnet pull and the stack is compressed by the difference in these two forces. During steady-state operation the three forces are always in equilibrium. That is

$$F_s - F_m = F_p \qquad (1)$$

where

F_s = spring force
F_m = magnet force
F_p = carbon-pile force

This general equation is true for any three forces as long as no motion is involved, but, to maintain a theoretically constant terminal voltage, the force values of F_m for all conditions of generator speed and load must be qualified. The magnet force F_m must be produced by a single steady-state value of voltage applied to the sensing circuit. The change in magnitude of F_m is caused by a change in the working air gap and not in the applied steady-state voltage. The transient voltage caused by a disturbance at the generator provides the force to cause the change in the value of the working air gap. If the qualifying condition for F_m is not fulfilled, the system voltage will vary to keep the three forces in equilibrium and the object of the design will not be fulfilled and the percentage regulation will be large.

As will be shown later in the discussion of Fig. 1, even from the theoretical approach the curves of the components only satisfy this theoretically perfect condition over part of the range of the carbon stack. This situation is a major limiting factor in maintaining a small percent regulation in wide speed-range generator systems.

Regulator Calibration

To construct a regulator to satisfy equation 1 and the qualifying statement which follows it, a certain sequence of steps must be adhered to which involves calibration and then final adjustment. To calibrate the regulator, the spring is assembled into the armature and calibrated against a standard test electromagnet. This is accomplished by a number of detail steps in which the spring loading and working air gap are varied. The armature is made to pick up and to release when the voltage applied to the test electromagnet is varied between certain predetermined limits. The action obtained is similar to the action of a relay armature. The calibrated armature is then assembled with a regulator housing, base assembly, and regulator electromagnet, and the electromagnet is calibrated to the armature by use of the core. The core adjustment determines the release voltage and the pickup voltage will then automatically be the same as the pickup voltage of the armature when it was on the test electromagnet. With this calibration complete, the proper force difference to operate the stack has been established. This force difference is shown graphically in the upper part of Fig. 9 by plotting the spring curve of Fig. 8 and the electromagnet characteristics of Fig. 7 together. The motion of the armature in the regulator with respect to the increase and decrease in voltage, without a stack installed, during the bench calibration can be determined graphically from this plot of curves 7 and 8 and is shown in the lower half of Fig. 9.

Referring to the upper half of Fig. 9, it can be seen that, when the voltage is zero, the magnet force is zero and the armature spring is at rest at point X_0. Projecting X_0 downward, point Y_0 is established at the intersection of the projected line and the voltage line V_0. Y_0 is the starting point of armature motion. The axes of voltage and armature motion are not conventional. The motion is with respect to the face of the magnet and not with respect to the point of the armature at rest. This method is used in Fig. 1 also. When the voltage is increased to V_1, the armature is attracted to the magnet until $F_s = F_m$. This condition is satisfied at the intersection of the two curves at point X_1. By projecting X_1 downward until it intersects with the voltage line V_1, the motion of the armature for the voltage V_1 is established at Y_1. This same procedure is repeated for all voltages V_0 through V_5. Points of equilibrium are obtained at all points except V_4 when the voltage is being increased, and V_3 when the voltage is being decreased. At the voltage V_4, the slopes of F_s and F_m become tangent at point X_4; and, if V_4 is increased a very slight amount, a value too small to show graphically the slopes of F_s and F_m become parallel at X_4' and then the slope of F_m becomes greater than the slope of F_s. When this occurs, the armature is suddenly snapped toward the magnet. The snap action stops at the first point where F_s is again equal to F_m and the slope of F_s is equal to the slope of F_m. This occurs at X_4''. The snap action is shown in the lower graph by the arrow (←). A similar condition occurs at the voltage V_3 when the voltage is being reduced and the armature is suddenly released or dropped out. These values of voltage at dropout and pickup are extremely important since they must fall within certain predetermined limits which signify that the desired force difference exists.

Regulator Adjustment

Now that the core position is fixed and the force difference required to operate the stack has been provided, the stack may be installed. The position of the stack with respect to the spring and magnet must be found, i.e., a final adjustment point must be determined. To do this, it is necessary to fix the generator operating conditions. The generator shall be operated at no load and at average rated speed requiring the stack to assume a resistance value of *N* ohms, see Fig. 6, to hold the voltage at its rated value. For simplicity, the length of the pile will be considered fixed so that the motion of the pile, as it is moved into place by the stack screw, will be the same as the motion in the air gap. This assumption also fixes the stack force. Maintaining a fixed stack length and force will introduce noticeable errors of a few thousandths in the stack-screw travel and a few ounces of the stack force only at the extremes of the adjustment procedure. These extremes occur only when the stack is first moved against the spring and when the spring is severely compressed by the stack and magnet beyond the point of normal regulator use.

By following the same general procedure, outlined for graphically determining the armature motion in Fig. 9, the variation in voltage of the system with respect to the rotation of the stack-adjusting screw may also be determined by referring to Fig. 1. The resulting curve is called the adjustment curve, as will be shown. At the instant the stack is

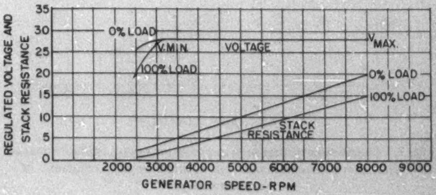


Fig. 10. Theoretical regulation characteristic. Generator speed rating 3,000 to 8,000 rpm

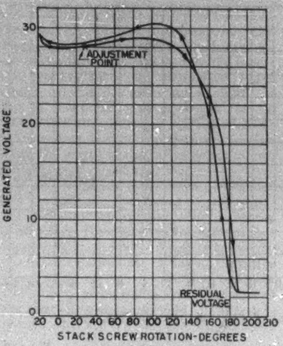


Fig. 11. Actual adjustment curve

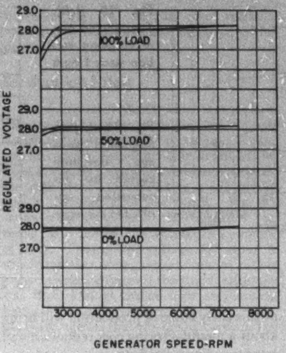


Fig. 12. Actual generator-regulator regulation curves

brought against the armature at Y_r, the stack force is zero, the stack resistance is infinite, and the voltage is the residual voltage V_r. The stack screw position is at point Z_r, the intersection of the projection of Y_r and the voltage line V_r. If the stack screw is rotated clockwise until the spring is compressed to the point Y_1, the stack resistance is immediately reduced to a finite value and the generator voltage increases. The value of this voltage can be determined by using equation 1 in the form $F_m = F_s - F_p$. The force of the spring is read from Fig. 1 as F_s' and the force of the pile has been established at N ohms as F_p'; therefore, by substituting into equation 1, F_m becomes the known value F_m'. The voltage required to produce the force F_m' is V_1. With V_1 known, the point Z_1 is established by projection. Following this procedure for a sufficient number of points will produce the curve in the lower half of Fig. 1.

To properly locate the pile with the stack screw, the qualifying statement of equation 1 must be met. If it is assumed that V_3 is the rated voltage which is to be maintained at a constant value, then the stack for this single speed and load condition must be placed so that, when equation 1 is solved, the force of the magnet is produced by the voltage V_3. This condition exists for a single load and speed at the three stack positions Y_2, Y_3, and Y_4.

Two of the positions may be eliminated by applying the qualifying statement following equation 1 at other generator speeds and loads. If the generator speed is increased, the stack will assume a higher value of resistance at a smaller force differential. The change in stack characteristics is shown in Fig. 6. Taking F_p'' as the stack force value for maximum rated speed at no load and plotting in Fig. 1 with respect to Y_2 yields point A_2 with the resultant F_m'' which does not lie on the curve V_3. A similar solution at Y_4 also does not produce a resultant magnet force which falls on the curve V_3. The solution at Y_3 is the only one of the three that satisfies the qualifying statement. This can be worked out similarly for the lower speeds with the same answer; however, at minimum rated speed and full load the characteristic of the stack will not satisfy the qualifying statement at any point and the system voltage must droop to produce equilibrium. The complete stack characteristic is plotted in Fig. 1 where the greatest matching of the stack curve and voltage curve occurs. It can now be seen that for best operation of the regulator the voltage curve produced by the stack screw can be used as an adjustment curve and that the correct point of adjustment of the stack with respect to the armature and magnet is at point Z_3 when the adjustment is made at average generator speed and zero load.

Theoretical Regulation

The regulating characteristic to be expected after the calibration and adjustment are complete and can be determined by reference to Fig. 1 which includes the characteristics of all components. Knowing for a given generator design the resistance variation required over its speed and load range, the voltage or voltages which will be generated are found at the intersection of the stack curve with the pull curves of the magnet for the particular stack resistances being considered. The resultant curves of voltage and stack resistance with respect to generator speed and load are shown in Fig. 10. The voltage regulation to be expected under these theoretical conditions without including the effect of hysteresis and friction is

$$\text{Per-cent regulation} = \frac{V_{\max} - V_{\min}}{V_{\max}} \times 100$$

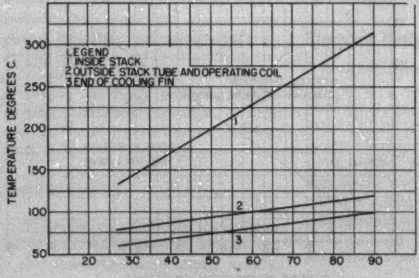


Fig. 13. Regulator temperatures at a room ambient of 25 degrees centigrade

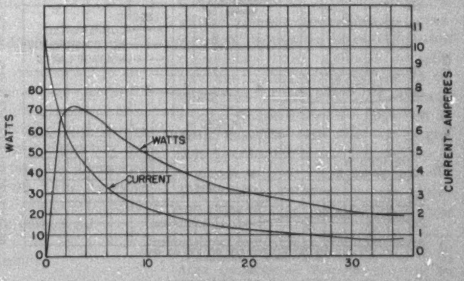


Fig. 14. Relation of stack current and stack watts with respect to stack resistance for a generator-regulator system

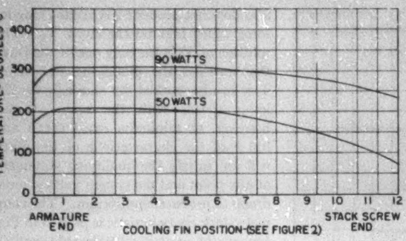
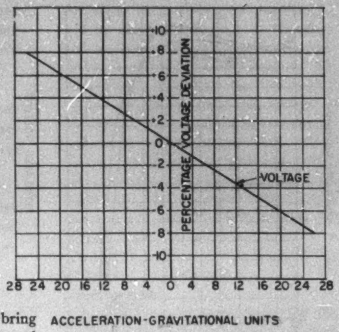


Fig. 15 (left). Temperature gradient along carbon stack length

Fig. 17 (right). Regulated voltage deviation caused by acceleration applied to the regulator



Conclusions

Through the graphical analysis of the component characteristics, as shown in Fig. 1, it has been shown that, with regulators designed to be stable throughout the adjustment range, there is a definite optimum point of adjustment on the adjustment curve. This single point is Z_3. For a given design, the point Z_3 bears a fixed relation to the bottom of the dip in the curve, therefore, the adjustment does not rely on the skill and judgment of highly trained personnel who are well versed in the trial and error method of adjustment. It does mean, however, that with the proper education the adjustment of regulators may become simplified and standardized.

During regulator assembly, before the stack is installed, it is desirable to check the calibration of the armature and magnet for two reasons. It causes the magnet core to be located in the correct position for final test and, by noting the amount of difference in voltage between armature pickup and release, it provides a means of weeding out those regulators which will not provide the proper force to operate the stack. Correcting the defective units at this early stage of production will bring about a manufacturing cost reduction and a more uniform product.

Appendix. Hysteresis, Friction, and Environmental Effects

From the practical standpoint, there are many interesting factors affecting the regulator operation which must be considered in the design for aircraft application. Some of the factors are interrelated and when they occur simultaneously they may produce a general effect which is slightly different than when they occur singularly. To best present a few of these basic factors, they will be considered singularly in most cases.

Hysteresis, Friction, and Curve-Matching Effects

The effect of hysteresis and friction on the adjustment curve can be seen by comparing the plot of test data, Fig. 11, with the theoretical curve; see Fig. 1. Since this loop exists, the question arises whether the upper or lower half of the curve ought to be used. The standard procedure which has existed for a number of years is to use the lower half.

Similarly, the actual regulation data taken on a generator-regulator system, Fig. 12, can be compared with the theoretical regulation curve; see Fig. 10. It will be noted that the basic shape of the theoretical curve shown in Fig. 10 is present but that the curves assume a definite slope and are not reversible. This condition is caused by three factors which were not included in the theoretical approach. These factors are hysteresis, friction, and the fact that perfect curve-matching on a production line basis is impossible. The regulating limits which apply in this case are 27.2 to 28.8 volts between 3,500 rpm and 8,000 rpm. The regulator performance is well within these limits which would make them seem very liberal, but this is a false impression since the regulator must also remain within these limits when the ambient temperature is varied between −55 and +71 degrees centigrade and when it is subjected to numerous other environmental conditions.

Temperature Effects

Fig. 13 shows temperatures at different parts of the regulator for various wattage dissipations and also indirectly shows the temperature gradient from the carbon stack to the end of the cooling fins at different wattage dissipations in the stack. The gradient at any wattage is found by drawing

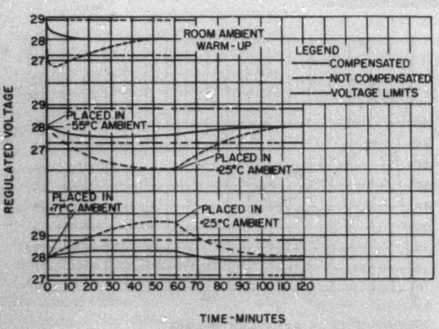


Fig. 16. Variation of regulated voltage caused by variation of the regulator ambient temperature

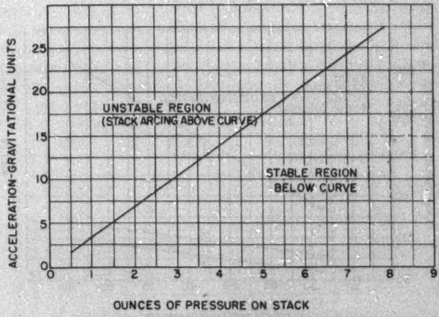


Fig. 18. A basic curve showing the stability of the regulator dependent upon the force under which the stack is operating

a vertical line through the wattage value and then reading off the temperatures. The wattage and stack current to which any regulator is subjected is governed by the generator requirements. The requirements of a typical 30-volt d-c generator is shown in Fig. 14. Another factor, shown in Fig. 15, which does not carry much importance but is interesting is the temperature variation along the length of the carbon stack at two different wattage dissipations.

If the regulator is constructed without any temperature compensation, the variation in voltage is quite large and is caused primarily by the change in the resistance of the operating winding. A regulator may be compensated a number of ways. Negative temperature coefficient resistors, bimetals, and variable permeability shunts are three of the most common means used. The results of tests with and without compensation are shown in Fig. 16. In this case, variable permeability shunt material was used in the magnetic circuit.

Acceleration Effect

There are two effects which take place with constant acceleration: one is a voltage deviation, and the other is arcing. These two factors are noticeable only in one of the three mutually perpendicular planes through the regulator. The sensitive plane is the one in parallel with the stack. Fig. 17 shows what happens to the voltage. The single curve shown is an average of a number of tests. During this voltage deviation, it was found that arcing would occur at rather low accelerating forces if the stack resistance was high, and at rather high accelerating forces if the stack resistance was low. This, of course, means that the arcing is related to the force on the stack. This basic relation is shown in Fig. 18. The slope is subject to variations with differences in the regulator design.

References

1. Carbon Pile Voltage Regulators for Aircraft, W. G. Neild. *AIEE Transactions (Electrical Engineering)*, vol. 63, Nov. 1944, pp. 839–43.

2. Analysis and Redesign of a Carbon Pile Regulator for Aircraft Generators, W. B. Kouwenhoven, G. J. Thaler. *AIEE Transactions*, vol. 67, pt. II, 1948, pp. 1197–1203.

3. Aircraft Carbon Pile Voltage Regulators, Fundamental and Design Improvements, B. O. Austin, H. H. C. Richards. *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 983–87.

4. Damping Device for a Carbon Pile Regulator, Bascum O. Austin. *United States Patent 2,408,188.*

High Altitude Aircraft Inverters

P. W. FRANKLIN
MEMBER AIEE

OF ALL rotating electric aircraft equipment, the inverter is probably one of the most difficult and most interesting items. A semi-independent power conversion unit, it presents a combination of design problems inherent to its components, together with additional ones, such as special vibration, ventilation, starting, and regulation problems. Some of the inherent problems are partially eased by its limited output such as, for instance, commutation; others, such as symmetry requirements, are aggravated. Thus, the inverter is a highly compacted motor generator set, regulated to very close limits, self-ventilated or, lately, mixed ventilated throughout a considerable range of altitudes and temperatures, vibrated with as high as 500 cycles per second (cps) equivalent to accelerations of 5 to 10 *G*'s, and exposed to shocks, fungus, humidity, frequent starting, etc. In the following, an attempt will be made to outline the design problems, discuss possible cooling and regulation methods, and anticipate future lines of development.

General Design

At present, it appears that the majority of inverters to be used in aircraft will center around two major ratings, namely, 750 and 2,500 volt-amperes, either 3- or single-phase, up to 0.9–0.75 lagging power factor. These ratings are built with speeds of either 8,000 or 12,000 rpm, the latter, of course, slightly lighter, but posing increased mechanical design problems. Shock and vibration isolation is either relegated to particularly sensitive parts, or isolation of the complete inverter is occasionally taken recourse to, which offers certain advantages. The control box, containing regulating equipment, filters and, if required, relays, is mounted integral with the inverter; but, if the airframe manufacturers wished to and would provide for leads and cooling, it could be separated from the inverter. Future designs will probably dispense with power control relays; inasmuch as the inverter cooling system is capable of including the control box, this feature, excluding extreme environmental conditions, could continue.

Electrical Design Problems

Design problems encountered in electric aircraft machines are essentially problems of adaptation of electrical design details to the possibilities and requirements offered by new insulation material, brushes, bearings, and dimensional and weight limitations. The theory required to design and predict performance is known sufficiently, and proved.

A-C Generator Design

The ultimate design limitation, as in any a-c aircraft generator, is the exciter field. Within a rather limited space it is required to combine sufficient space for the magnetic steel, for a flux, often defined by an imposed overload condition, a low reactance damper winding, that is capable of carrying the negative sequence current of the single-phase rating, space for the d-c excitation winding with reasonable current densities, all structural design details required to hold the winding in place, a geometrical arrangement inimical to high pole-to-pole leakage flux, and, last but surely not least, space for ducts carrying the cooling air as close as possible to heat-producing and heat-sensitive parts. The d-c winding also has to be laid out with regard to easy installation and with consideration of the very high winding temperature, to be expected at maximum altitude, or possibly at highest sea level temperature. This temperature as, for instance, measured by resistance, today approaches 220 C.

It is well known that the output of any electric machine is proportional to the speed, D^2L of the armature and the product $A_s \times B_g \times WF$. Inasmuch as wave shape requirements of the line-to-neutral voltage require a two-thirds stator winding pitch, the winding factor is limited to approximately 0.83. Offhand, the split-up of the product $A_s \times B_g$ is up to the designer and the quotient B_g/A_s indicates whether he has chosen a "copper" or an "iron" machine. The rotor has to provide space for the air-gap flux and copper section required to magnetize the machine and overcome armature reaction. From this point of view, the available space is

---

Paper 54-415, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 30, 1954; made available for printing September 20, 1954.

P. W. Franklin is with the Leland Electric Company, Dayton, Ohio.

Lists of Symbols

a = height of coil, inches
c = dimension of bare rectangular conductor, inches
c_h = specific heat, joules per pound, degrees centigrade (C)
d = thickness of coil, inches
g = gravitational constant, 32.2 feet per second^2
h_a = average heat transfer constant, watts per inch^2, C
i = current density, amperes per inch^2
k = composite thermal conductivity, watts × inches per C, equations 15 and 17
k_i = thermal conductivity of conductor insulation and varnish, watts × inches per C
m = correction factor for additional iron losses resulting from punching, burrs, nonuniform flux distribution, normally 1.5 to 2.2
n = dimension of bare rectangular conductor, inches
Δp = pressure differential, inches water
q = specific loss per volume unit, watts per inch^3
r_t = specific resistivity of copper $= 1.724 \times 10^{-6}(1+(t\text{-}20)3.93 \times 10^{-3})$ ohm-centimeters (cm)
s = specific weight of air, pounds per foot^3
s_F = specific weight of iron, 7/6 to 7.8 grams per cm^3
t_A = ambient temperature, local air temperature, C
t_F = gradient across film, C
t_i = gradient across insulation, C
t_c = gradient within coil, C
v_{10} = specific iron losses, determined by Epstein test, at rated frequency and 10,000 gausses, watts per pound
v_t = air velocity, feet per minute
A = equivalent orifice, inches^2
A_s = stator ampere conductors per inch
B = flux density, lines per inch^2
B_g = air gap density, lines per inch^2
C = dimension of insulated rectangular conductor, inches
D = diameter of bare conductor, inches
D' = diameter of insulated conductor, inches
D_0 = outside diameter of stator lamination, inches
F = free flow area, inches^2
F_1 = free flow area, d-c end of inverter, inches^2
F_2 = free flow area, a-c end of inverter, inches^2
F_t = individual free flow area, feet^2
N = dimension of insulated rectangular conductor, inches
P = weight of body, pounds
Q = air mass flow, pounds per minute
S_0 = surface, inches^2
T = time constant, seconds
WF = winding factor
W = loss, watts
W_0 = loss at cold reference temperature, watts
$\alpha = 3.93 \times 10^{-3}$
γ, β = constants, see equations 2 and 3
ξ_t = restriction coefficient
η = correction factor
ϵ = copper fill factor, see equations 14 and 16
ν = geometrical correction factor
τ = pole pitch, inches
δ = air gap, inches

---

utilized better if filled with copper than with iron; hence, for low-temperature machines and those without particular overload and symmetry requirements, a machine with a low B_g and high A_s will usually use less space or give a higher output for a given volume. With the advent of extreme temperature machines and the requirement of high sustained short-circuit currents, the tendency to be recommended is toward the iron machine with relatively little copper. Aside from the fact that such a design results in a machine with lower reactances, less armature reaction and a less distorted pole flux under load, the copper losses are lower and the machine is cooler. Also, the change in field current between *NL* and *FL* is decreased. Because of the relative increase in iron cross section, such a machine has a tendency to be heavier, but this effect can be minimized by suitably detailed design. A point of interest is the proper choice of the pole arc. As it increases, the ratio of useful flux (sinusoidal) to actual flux (essentially square-shaped) decreases; furthermore, the space in the rotor available for copper decreases. Consequently, there is a definite optimum pole arc which varies from case to case, but of which the designer accustomed to calculate details is well aware. An important factor that must be considered very early is the method of how to insert the rotor windings. Normally, integral rotor laminations are preferred, but this will result in random coils with a relatively poor space factor. Detachable pole components with precision-wound inserted coils add mechanical complications. A 4-pole structure will permit precision-wound coils with an integral rotor punching but with a certain loss of available winding space.

The damper winding is a most essential component in the case of single-phase operation, and can be considered as being connected in series with the stator winding. For good damping, and elimination of the negative-sequence third harmonic (which is of an entirely different nature from the pole flux distortion third harmonic and cannot be eliminated by pitching) continuous end rings are essential. The writer believes that present and future load and temperature conditions prohibit the use of solder in stationary and rotating parts and, in view of the nature of the damper windings, strongly recommends brazing or any other high-temperature connection. Even if stabilized temperatures should permit the use of solder, the inherent weakness of a soldered joint towards short-timed but high-current overloads, which otherwise may well be accepted by other parts of the winding, limits its use. This leads to another complication, namely, the need to braze continuous end rings after the d-c coils have been assembled, without injury to the insulation.

Of all the electrical components of the inverter, the exciter is usually the hottest and most sensitive to temperature, and it may be well to discuss the temperature problem here. It is also the component where the changes in resistance attributable to temperature is most critical, since at very high temperatures the exciter resistance may prevent the a-c voltage regulator from forcing sufficient current through the winding, and thus present an output limitation independent of insulation life or electromagnetic parameters. The temperature, as such, is limited by the available insulation.

Insulation life is essentially a function of the time-temperature integral and consequently a distinction must be made between required average life at average conditions and minimum life required at worst conditions. To put it plainly: present inverter design, and undoubtedly also all electric aircraft machinery design, has become so marginal, in view of the greatly increased environmental conditions, that the whole problem is ripe for a thorough review. It is, for instance, quite feasible to wind a 35,000-foot machine, operating at −10 C with heavy Formvar in the exciter, and to obtain acceptable life. At 50,000 and 65,000 feet with +20 C air temperature, this condition has changed radically. Unless other design proportions and parameters are sacrificed, hot-spot temperatures of up to 250 C and more must be expected in the exciter. Present insulation and varnish is quite capable of withstanding this temperature for a very reasonable period, but it is doubtful if the ambiguous general statement would apply, which is so often quoted in present specifications, that the machine should prove life for 1,000 hours in "any combination of altitude and temperature," thus including 1,000 hours' life at worst condition. This is neither economical nor necessary. Present margins in design require a careful balance between logistical factors, overhauling periods, initial cost, simplicity of production and, of course, weight and bulk required. Finally, the fact must be recognized that out of 1,000 hours' over-all life, a combination of worst altitude, load, and environmental temperature occurs actually only for a relatively short fraction of time. Present informal tentative figures indicate a ratio of total time at average conditions, including ground duty, to time operated at worst conditions of at least 4 to 1, with this ratio increasing as planes fly higher and faster. It is, of course, quite reasonable to ask for long life if this can be obtained at reasonable production cost, as well as weightwise. However, if these factors increase sharply when plotted against life, a revision of the basic idea must take place. As a first guide of what may be acceptable for life under worst conditions, three indications exist, namely, the total cost per hour plane flight versus the cost of the particular equipment under consideration, the total life of the plane, and finally present and future plane overhaul periods. These considerations form a subject for a separate paper and will not be pursued here. However, the writer wishes to state that proved life under a combination of worst conditions of 200 hours is entirely feasible, and is sufficient.

Present insulation problems are in a state of flux, and constant and gradual improvement in maximum permissible temperatures, space factors, and the mechanical strength of varnishes is observed.

The slip rings do not pose any particular problems, provided they are properly designed for high-temperature operation; this requires minimizing of all insulation and a proper choice of ring material and brush grades. It should be noted that the occasional requirement of successful operation from 17 to 29 volts, regarding the a-c generator, is basically a problem of permissible losses within the voltage control. Otherwise, the generator design is uninfluenced by this requirement, except that, at the expense of weight and bulk, these regulator losses may be reduced. This is not the case with the d-c motor.

The a-c stator is much easier to design than the exciter. This is owing to the stationary nature of the windings and the fact that the total current volume in the stator compared with that of the exciter winding is only 60 to 75 per cent (%) and is located in an expanding geometry. The stator is usually designed for three phases, with six leads, but must be capable, when connected in delta, to carry full single-phase load. This causes a doubling of copper losses for single phase. Also, single-phase leakage reactance, including the damper cage leakage reactance, is higher than that for 3-phase load. Consequently, the machine must be designed basically for single-phase load, and will have higher losses and about 5 to 10 per cent % higher field current values than for 3-phase load. Single-phase efficiency will drop accordingly. It should be noted that single-phase damper losses may approach in magnitude the primary copper losses.

Wave shape is of the essence with regard to line-to-line and line-to-neutral voltages. Pole flux distortion harmonics may be controlled by using a two-thirds coil pitch; slot harmonics are reduced by skewing and choice of the number of slots. The negative sequence harmonics are mostly a function of the leakage reactance and saliency of the damper winding; hence, round bars with slot openings and continuous pole rings are highly recommendable.

The stator magnetic circuit is conventional; iron densities should be chosen entirely with regard to required exciter ampere turns and not with regard to iron losses, since the stator is usually easy to cool and less sensitive to high temperatures. High-grade silicone steels in thin laminations, annealed after punching, are undoubtedly standard procedures. All stator winding connections should be brazed.

In one detail the a-c generator varies from its commercial counterpart, and this applies to a-c as well as d-c machines. Because of weight restrictions and the necessity to compress more power into available space, the ratio of air gap to pole pitch is made much smaller than in commercial equipment. Here, this ratio is usually determined by pole flux distortion and manufacturing techniques. In aircraft machines, the air gap is chosen primarily with regard to reduced exciter or shunt field turns, limited mostly by mechanical considerations. A representative ratio is $A_s\tau/B\delta_g$. In aircraft machinery this ratio is about four times as large as in commercial machines and the consequences are serious. First, surface and pulsation losses increase considerably; second, the pole flux is distorted considerably owing to slot openings and armature reaction, and negative flux areas are not uncommon. This further increases the losses, particularly tooth volume losses. The distortion is so strong that compensation by pole chamferring becomes almost impractical, and the only way to filter out harmonics is by pitching, skewing and, occasionally, fractional slot per pole windings. The gain in pole flux along one pole edge is more than counterbalanced by the loss along the other, requiring an increase in total exciter ampere turns. This is a phenomenon commonly encountered in d-c machines, and it occurs also in a-c machines. Finally, the pole coefficient as, for instance, the ratio of fictitious first harmonic amplitude to actual average air-gap density, direct and quadrature coefficients, become now a function of saturation, and vary more or less with it. All these factors, strongly pronounced in aircraft machines, make the exact prediction of performance rather difficult.

Typical inverter a-c generator design data are as follows:

Stator copper densities (single-phase connection) 6–8,000 amperes per inch^2
Air-gap density (sinusoidal peak)—B_g = 30,000 to 42,000 lines per inch^2.
Stator ampere conductors per inch—A_s = 500 to 800.
Rotor copper densities—7,000 to 8,000 amperes per inch^2 or less.

D-C DESIGN

The d-c motor operates at constant speed with a nominal voltage of 27.5 volts and with armature currents up to approximately 180 amperes for the largest ratings. Commutation is consequently easier than in conventional aircraft d-c generators. As such, the commutation problem is the same as in d-c generators and can be approached by the same methods used there. Of importance are a proper single clearance, low-reactance voltages, which usually can be obtained, proper choice of coil pitch and slot numbers and a good mechanical commutator and brush holder design.

The design limitations are now as follows: A motor is required to operate continuously over a voltage range of either 26 to 29 volts or, occasionally, from 17 to 29 volts, the latter for limited periods. The starting current is to be limited to 1,000% rated current, which either limits efficiency or requires special relays and a slightly more complicated winding. Depending upon the decision of the designer, a more or less limited series winding is included, part of it cut out after starting or left in permanently with ensuing shunt field regulation problems. In view of the aforementioned pole flux distortion and, particularly, if the operating voltage range is wide, a compensation winding is necessary, which should match the armature reaction fairly closely. In view of the relatively small commutation reactance voltage (normally below 1.2 volts) most of the commutation troubles are of a mechanical nature, and it is possible to omit interpoles and still obtain acceptable commutation. It should be noted that in aircraft machines the additional tooth tip reactance, caused by the interpole shoe, is of considerable magnitude, and that the addition of an interpole inherently increases this fictitious reactance voltage. The omission of interpoles, and the use of compensating windings only, improves efficiency and, most important, assists cooling by providing additional air passages.

In a similar manner as in the a-c machine, a machine with strong field is far preferable since it results in better commutation and cooler armatures; and since the shunt field geometry is expanding, less trouble is encountered with shunt field space requirements.

Armature windings, depending upon size and manufacturer traditions, are either wave windings or lap windings, well equalized, and both have given good results. Either solid conductors or magnet wire may be used. Magnet wire, slot insulation and varnishes suitable for hot-spot temperatures up to 300 C for limited periods are available or appear to be so within the next few years. All connections should be brazed and the use of solder as a balancing aid should be avoided. Commutator design practices similar to d-c generators, with regard to both material and geometry, are used. In general, the same range of values prevails for A_s, B_g, and current densities as for the a-c machine.

It is well known that, in aircraft d-c machines, because of the high values of A_s and relatively wide brushes, the armature reaction location depends very much upon the center of current collection under the brush, the effective neutral. As such, this location will change as a result of variation of contact conditions, caused either by varying filming, changes of the friction coefficients, and humidity, altitude effects, etc. In voltage regulated d-c generators, this shift of neutral may be observed as it manifests itself in varying shunt field currents; in inverter motors, the same phenomenon, particularly at altitude, causes sudden jumps in armature currents; and the degree of observed steadiness is a good indication for the particular brush condition. This phenomenon is often aggravated by the transient characteristics of the d-c power supply and the resistance of the power lines.

The relatively large effect of the location of the neutral requires fairly close adjustments of the brush holder rigging. Such an adjustment can be performed easily in the field by semiskilled personnel, requiring 115-volt 60-cycle power, 100 to 150-watt light bulbs, and a small a-c voltmeter. The bulbs are connected in series with the armature across 115-volt alternating current and the brushes are shifted until the shunt field voltage reads a minimum. Provided the brushes are formed, but not necessarily filmed, this position is then the neutral position. In view of the ease of this adjustment, the writer feels that the present general military specification requiring machines not only with nonadjustable neutrals, but even completely interchangeable without further adjustment, may be unnecessary. It imposes a considerable number of difficulties upon the manufacturer and, to be quite frank, cannot be complied with on machinery with high design parameters. Regardless of how well production is set up, the accumulated tolerances will not permit a uniform product in this respect. For instance, during the pressing-on operation, a commutator may be guided until its bore bites the shaft; from then on, the commutator slides down and, if it chooses to skew gradually in descending, it cannot be prevented from doing so. The writer feels that the present tendency in d-c motors under development, namely, the tendency to permit individual adjustment, is much more adequate and acceptable to all concerned, provided that, after disassembly and subsequent reassembly, the original position may be restored.

REGULATORS

Since this component is one of the most proprietory ones, only a systematic discussion will be presented here. Basically, the field power may be taken either from the a-c end or the d-c end. The former is the more appealing inasmuch as a complete static regulator with magnetic amplifiers could be used, and thus several of the disadvantages of the methods to be outlined are avoided. However, new difficulties are introduced. First, the output of the inverter has to be increased approximately 15 to 20%, with a consequent increase of the motor. Self excitation is problematical; a sustained a-c short-circuit current may be obtained only by the help of additional relays. Selenium rectifiers must be used at the present time, and their operation in an ambient of 120 C is again problematical, at least for the time being.

A more logical approach to excitation is, of course, the use of the original d-c power. The sensing device may consist of resonant circuits for both voltage and frequency, cold cathode tubes, thermal diodes, or a magnetic-mechanical device, counterbalanced by a special spring. Amplification may be obtained either electronically, by magnetic amplifiers, or within the magnetic structure of the power-controlling device which controls the field currents, and which is either a carbon pile device or a finger-type regulator. In the first case, the limiting factor is watt dissipation of the pile, under worst environmental conditions; in the second, the maximum power interrupting capacity per finger. Both exert a serious restriction upon machine design. Points to be considered with regard to the choice of a good regulator are:

1. Accuracy.
2. Reliability and maintenance in the field.
3. Life under unfavorable environmental conditions.
4. Influence upon system characteristics.

With regard to accuracy, an electronic sensing device, operating at very low grid power levels, is probably the most accurate. Reliability and serviceability in the field, however, tend to favor devices with lower accuracy. Resistance to high temperatures discriminates at this moment against selenium rectifiers, unless special cooling conditions are integrated into the design. At this moment, a trend is noticeable toward the carbon pile, either trimmed or directly steered by sensing devices, operating essentially with resonant circuits or temperature independent nonlinear circuit elements, amplified by magnetic amplifiers of low-level power. In all cases, where a-c power is being used for amplification purposes, provisions must be made for sustaining a sufficiently large short-circuit current, usually above 200% rated current for a symmetrical 3-phase short-circuit condition. All regulation systems should, of course, adhere to the fail-safe principle.

Air Flow

At present, inverters are required to supply their own air flow and are rated up to altitudes of 50,000 to 65,000 feet, with ambients as high as +20 C. Lately, however, a trend has been started either to superpose or replace self-ventilation by blast cooling, similar to that of other rotating electrical equipment. In the prediction of air flow performance of inverters, two concepts have been found to be verified fairly well by experience and are of considerable assistance:

1. The fan or fans may be considered as a constant volume (constant displacement) device. Consequently, the pressure created is proportional to the specific weight of air, as is the fan horsepower.
2. The inverter, at constant speed, may be considered as an equivalent orifice.

For high altitude fans, fan pressure measured at sea level will usually range from 10- to 16-inch water; and the equivalent orifice of a 2,500-volt-ampere inverter will be of the magnitude of 2- to 2.5-inches² (assuming an orifice coefficient of unity).

The use of a constant volume fan leads to a dilemma. If it is designed to supply sufficient air mass flow at high altitude, then at sea level the mass flow will necessarily be approximately ten to fifteen times as much, with a consequent increase in fan horsepower and decrease in efficiency. However, this permits the inverter to operate in rather high ambient temperatures at sea level, up to 120 C, without any particular detrimental effects. Airflow calculations are made on the assumption of incompressible flow; for a refined calculation, a differentiation between the density of the air within the fan and the machine proper could be made. Inasmuch as the assumption of incompressibility holds for changes of the ambient pressure of up to 10%, and the fan pressure versus ambient pressure is approximately constant, this assumption is fairly well justified even at high altitudes, which does not necessarily hold for blast cooling.

Calculations of air flow within the machine are very difficult to perform for a new design, and even for one where test results are available the duplication of such results by well-known methods is not always easy. Particular difficulties exist in predicting entrance shock losses into rotating grids like commutator risers or pole wheels.

The following air passages usually exist:

1. Ventilation around the commutator, sometimes with special baffling around the brushes.
2. Air passage outside the d-c stator.
3. Air passage within the d-c stator, interpolar spaces, etc.
4. Air passages through commutator risers and below armature core.
5. Air passages outside of a-c stators.
6. Air passage within pole wheel (a-c).
7. Air passage around slip rings.
8. Entrance and exit.
9. Air passage within control box.

One or several fans may be used and placed either in the air intake where cooler and denser air exists, resulting in larger mass flow or in the exit, which has the advantage that the additional temperature rise of the air as caused by the fan is not being communicated to the machine. A fan, placed at the air entrance, is capable of providing pressure for control box ventilation, and is therefore preferable. The previously enumerated air passages can now be connected in two principal ways, by the duplex system or by the simplex system.

Duplex System

By the duplex system, air is taken from both ends and exhausted in the center of the machine. Control box air is bled from the machine from one end, preferably before the brushes have contaminated the air. Both ends do not necessarily carry the same mass flow. Both bearings and brush sections receive cool air; furthermore, because of the parabolic nature of air-flow resistance, the total fan horsepower for the same amount of total cubic feet per minute is slightly less than for simplex airflow. However, this system is not as easily fitted into the airframe as the simplex system.

Simplex System

By the simplex system, air is taken in at the d-c end, blown through the complete machine, and exhausted at the a-c end. The machine fits well into most airframe manufacturers' ventilation ducting; the d-c brushes receive a considerable amount of cool air, which, if well ducted, assists in the brush cooling. If one fan is used, certain structural simplifications are obtained. The a-c exciter receives hot air, but with increased mass flow, and it can be shown mathematically that its resulting temperature rise does not necessarily exceed the temperatures obtained with the duplex system with the same total air mass flow. The slip rings are located in high air temperatures, but since the brush densities and resulting slip ring losses are low, this does not necessarily lead to major difficulties. However, the a-c bearing is also in the path of the hot air and, since present-day greases are limited in temperature range, it forms a definite design limitation. Furthermore, since minor brush dusting will always take place at higher altitudes and the cooling air will carry traces of brush dust, the a-c bearing is thus exposed to contamination. Sometimes, it is possible to by-pass cool air around the machine and bring it into contact with the a-c bearing, thus lowering its temperature.

It is possible to follow up the air flow numerically, although it must always be implemented by experimental verification and certain empirical data. With a known cubic-foot-per-minute airflow, the mass flow at various altitudes can be fairly well predicted, provided the air density is known. The writer was able to verify this surprisingly closely by measuring air-in and air-out temperatures at various altitudes and comparing this rise with the predicted rise, obtained from measured total losses and assumed air mass flow. The total fan pressure may be fairly well measured by a single-end Pitot tube, held to the end of the fan, and adjusted so as to give maximum readings to obtain a combination of pressure and velocity head. This pressure was found to be approximately up to 80% of the pressure calculated with the use of Euler's equations. The fan horsepower, appearing as an increase of d-c input, may be calculated with a total fan efficiency of 45 to 60%. The total equivalent orifice, presented by the machine, could theoretically be predicted by assuming various air paths in parallel and series connection. However, the writer found that equivalent orifice data, measured and evaluated as a proportion of the total ducting area in the center of the machine, resulted in a surprisingly constant percentage, varying from 40 to 60%.

Several equations dealing with air flow are given in the Appendix. Their main purpose is, at this time, to provide a systematic approach for analysis and recalculation for cases where test results are known. Such results, analyzed and condensed into proper constants, can then be employed to predict future similar designs.

The problem of superposition of external air blast cooling upon an inherent inverter fan has recently come under consideration. For a given volumetric flow, velocity relations will be established within the air intake, fan, and diffuser, if used. At the optimum point, the shock losses, caused by mismatching of tangential speed, approach, or exit air speed and blade angles, will be at a minimum. Above and below this point, the fan efficiency will drop. If, as a result of added pressure, a larger volume is forced through the fan, the losses will increase, and if the added pressure is such that the volume is beyond the capabilities of the fan, the flow pattern within it will be disturbed to a degree that pumping action is replaced by mere turbulence. The writer understands that such cases have already been observed. The amount of permissible pressure superposition therefore has to be investigated for each fan and air density condition, and depends primarily upon the pressure-weight flow characteristics of the fan.

Temperature Distribution

The problem of calculating the temperature distribution within electric machinery is almost as old as electric machinery itself. Attempts to approach the temperature field in machines by one or several second-degree linear differential equations, coupled by boundary conditions with an additional first-degree differential equation for the cooling medium, also coupled to the aforementioned system, are well known and probably regain practical importance with the increased use of electromechanical calculation devices.[1,2,3] However, their accuracy is limited by the accuracy predicting the flow of the cooling medium and the accurate knowledge of heat transfer as well as material constants. In view of the fact that these constants vary within fairly wide ranges for machines of the same type, because of variations in the varnish, the thickness of varnish layers, variations in slot cell material, variations in the size of the usually small cooling ducts, etc., detailed calculations for routine purposes are of doubtful use. The great usefulness of such detailed calculations for very large machines is undisputed where boundary cooling conditions are easier predictable and for academic and basic work.

Yet, for several reasons, a fairly simple calculation method is necessary for aircraft machines. First, there is the problem of rerating, which is closely connected to temperature rise. Second, a simplified approach to this problem, even if it is not very accurate, helps in establishing a systematic approach in analyzing the problem and evaluating and condensing test results into representative parameters.

The inverter consists essentially of four groups of equipment, some of which are semi-independent as to temperature rise. They are: 1. the control box; 2. bearings; 3. windings; and 4. brushes.

Control Box

It is doubtful if the control box of a fairly large inverter could be cooled entirely by natural convection for altitudes of 50 to 65,000 feet. Hence, it would have to be placed either in a separately cooled compartment or receive air from the cooling circuit of the inverter. As such, it would obtain environmental air, raised slightly in temperature because of fan work. This may not be too critical at higher altitudes with +20 C, but becomes very marginal at sea level when air temperatures approach 120 C. In such a case, present-day selenium rectifiers reach their limitations, and mounting these rectifiers ahead of the fan air intake, has been considered, to obtain air as close to ambient as possible. Better rectifiers may help solve this problem. The relative positions of components within the control box are also critical, and heat-dissipating parts must be carefully shielded from those which are sensitive to temperatures. With the exception of selenium rectifiers, and as long as weight and bulk are not too much of the essence, all other components as tubes and capacitors are available for the aforementioned very high temperatures and reasonable life.

Bearings

Bearings have lately become one of the major temperature limitations in rotating electric aircraft machinery. Essentially, the problem is a grease problem and, as such, one of obtaining a grease with a wide temperature range, from −55 C to +160-to-170 C. It is not yet a metallurgical nor a shielding problem. It has been found that the bearing temperature, as measured on the outside race, differs for a certain amount from the environmental air temperature, this differential being first a proprietary design figure depending upon cooling surface, and also a function of air density and mass flow. At sea level the writer found values from 5 to 15 C, at 50,000 feet from 20 to as high as 40 to 45 C, depending upon the exposure of the bearing. As to bearings, the duplex airflow is unquestionably a much better design, since cold air is blown over the bearings. For such a design, present-day greases, which have acceptable performance characteristics from −55 to +125 C, are available. Unfortunately, the duplex design does not always suit the geometry of the installation, and it is the otherwise highly desirable simplex airflow, which is badly limited by the temperature of the exhaust bearing. There again, a single-fan installation is preferable to a double-fan type, since the exhaust bearing can be exposed better to air. Occasionally, cool air may be guided directly to the exhaust bearing through several by-passes within the machine. On the whole, the situation is by no means satisfactory, and the writer feels that little can be done from the design aspect within present required performance characteristics, to compensate for the absence of a proper wide range grease.

By assuming a typical temperature differential between bearing race and air, maxumum permissible bearing temperature with given inverter losses and ambient, the minimum mass flow of air can be easily calculated. This will be one of the limiting conditions governing the rerating question.

Winding Temperatures

Two aspects exist, one suited to the rerating problem, the other to predicting approximate hot-spot temperatures. The latter will be dealt with first.

APPROXIMATION OF HOT-SPOT TEMPERATURES

It is well known that the life of insulation is a function of the temperature-time integral, and that the relation between temperature and available life is essentially an exponential one. These curves are already available from insulation suppliers and are often based on fairly realistic tests. It will be noted that these test results stray fairly wide and are, of course, among other details, dependent upon imposed mechanical stresses and also the combination of varnish, layer insulation and wire coating. The wide variation in material characteristics as a function of manufacturing processes, and the unavoidable variations owing to the winding process proper, make the concept of a temperature danger zone rather than that of a dangerous hot-spot temperature appear reasonable. Also, the often quoted law that a certain small temperature increase halves the expected life seems to be rather an indication of a tendency than a more or less fixed law.

The hot-spot temperature consists of the following gradients above ambient:

1. Temperature rise in air up to location close to hot spot.
2. Gradient between air and cooling surface adjacent to hot spot.
3. Gradient within various layers and space separating hot-spot location from surface.

The temperature rise of the air up to hot-spot location can be fairly well approximated, provided air flow conditions are known. The location of the hot spot can be fairly well estimated under full-load conditions and for conventional loss ratios. For stators and rotors cooled by axial air flow, the writer usually places the probable hot spot within the slots, about one-fourth core length away from the hot end. Because of the relatively flat temperature field distribution, this is usually accurate enough.

The gradient between air and surface can be fairly closely approximated by well-known equations. It should be noted that the heat transfer constants encountered in electric machinery are usually larger than those calculated based on either simplified laboratory duct tests or theoretical calculations, because of the added turbulence and bouncing of air between opposite walls of a duct. Various writers have indicated this ratio to be as high as 1.5 to 2.2.

The gradient within the body may be handled by using the theory of heat flow in a 2-dimensionally infinite plate with inherent volumetric loss, superposed with linear heat flow. These results have been known for a long time, and any novelty involved consists essentially in the magnitude of gradients and materials constants.

Basic heat flow calculations in electrical rotating machinery deal with a linearized conductor, representing a slot wire bundle, insulated, surrounded by a cylinder of iron laminations, which in turn is cooled by air. This configuration is joined on both ends by another wire bundle, insulated and cooled directly by air. This composite body, simulating one slot with two half-winding extensions, and assuming tangential symmetry around the armature (stator and rotor) is described mathematically by three second-degree differential equations, with a linear equation for the iron (in approximation) and a first-degree linear differential equation for the axial air flow. Sufficient boundary equations may be established.[1,2] If the conditions in the area of the hot spot are investigated, it is found that in this section the axial temperature distribution within the wire bundle is fairly constant. From this it is deduced that within this particular section only a small percentage of the copper losses is removed axially, and that most of the copper losses, say 85 to 90%, flow radially through the slot insulation into the iron and toward the iron surface. Axial heat flow across laminations is much smaller than radial flow and may be neglected for this application. This permits linearization and thus relatively simple equations may be obtained. With these equations properly used the writer has recalculated hot-spot temperatures up to 400 C for high temperature d-c machines and has obtained quite good correlations, some of them of course by benefit of the privilege of afterthought. Such may be used at least as a schematic approach for certain types of machinery and, after having been correlated with tests, will serve for further predictions. The principles of heat flow within a very large plate with volumetric losses are well known; however, they have to be adapted to the pecularities of electrical coils with either precision-wound coils or random windings. Applied to aircraft machinery, the well-known fact is verified that, when current densities of say 8,000 amperes per inch^2 are used, and average winding temperatures of 200 C are considered, the maximum permissible distance of expected hot-spot location-to-coil surface is limited to 3/8 to 1/2 inch.

Basically, two major heat flow patterns exist. In aircraft machinery, the first one is that of a slot bundle, surrounded by heat-producing iron, and accepting some losses across an air gap; the total heat flowing either to cooling ducts within the iron or to the iron back surface. The simplified model consists of a sector equivalent to one slot pitch, say, 1 inch long, with radial heat flow only. A known amount of loss enters the iron from the slot, in addition to an assumed fraction of the losses of the opposite member, surface losses, volumetric losses within the tooth section and finally, losses within the core. These losses known, and the geometry available, the gradients can be fairly well approximated. If the slot contains two conductors, the gradient within the slot is usually not large; however, gradients of up to 20 to 30 C were occasionally measured by thermocouples in stator slots; which, incidentally, vary approximately with the square of the inductor current.

The second heat flow pattern is that of a field coil mounted on a salient pole of an a-c generator. Again, a section is used limited to, say, 1-inch length axially; it is assumed that 10 to 30% of all copper losses are removed axially, and that the rest have to travel to the surface of the coil. Two possibilities exist: either the pole mass as such is so hot that no heat travels from the coil into the pole, in which case this can be approximated by a symmetrical plate of twice the coil thickness, or that the heat flows symmetrically to both sides of the coil, equally well into the pole and the air stream. Comparison with careful tests will indicate the proper approach. All of the basic equations are outlined in the Appendix.

Very little has been published with regard to heat transfer across air gaps. For a first approximation in the case of stationary or rotating salient poles, it may be assumed that relatively fresh air is swept across the pole surfaces. If it is assumed that the air velocity approaches zero speed at the stationary member, and rotating speed at the other, the average velocity of the air is half the circumferential speed; furthermore, in the case of the usually small air gaps used in aircraft machines (0.010 to 0.030 inch), the axial velocity is small. Since the air density may be approximated, the Reynolds figure may be calculated and is usually be found to well within the laminar region. This, owing to slot openings, etc., is obviously not the actual case, and considerable turbulence must exist. Therefore, a multiple of the theoretical Reynolds figure can be assumed for the first approximations. An exchange of opinions and publication of test results, if extant, would be highly desirable with regard to this point.

APPROXIMATIONS FOR RERATING PROBLEMS

Aircraft electric equipment is designed to provide output under certain limiting environmental conditions and is tested for compliance. The airframe manufacturer, using this equipment under a multiplicity of conditions, has to estimate how the permissible output will change with variations of environmental conditions or, for a given output and certain environmental conditions, has to determine the coolant supply necessary for satisfactory operation. The following variables are involved: losses in machine, mass flow and temperature of cooling air, as entering the machine, and permissible maximum machine temperatures. Each component of the machine will have its own relationship for these variables, and individual limiting boundary conditions will have to be superposed for the over-all limitation of the complete set to be arrived at. One of these limitations was furnished by the exhaust bearings and the use of selenium rectifiers; in the following, only the windings proper of the power section will be discussed. It should be noted that in the case of a simplex machine, the bearing temperatures depend upon the machine losses.

To avoid expensive testing, which actually should be conducted on a statistical basis, since one machine is not necessarily representative, recourse is taken to a simplified mathematical model from which the approximate boundaries are derived to be spot-checked by actual testing. For this purpose, the inverter, and all air-cooled equipment, has already been represented by an equivalent orifice, either blast cooled or constant-volume cooled; in addition to this, it is now also represented as a simple heat exchanger with one representative wall temperature. This second approximation is by no means a close one, yet it is workable and must be accepted in the absence of anything better. This concept, first developed by Dr. E. O. Nauman states that an output limitation is reached if the fictitious wall temperature reaches a certain degree. This temperature may be calculated by adding to the air-in temperature the complete or a fraction of the total temperature rise of the air, based on known total losses and mass flow; to this the temperature differential between average wall temperature to average or final air temperature is added. The last term depends upon known losses, an empirically found lumped product of average surfaces times heat transfer constant per C, and the 0.8th power of the mass flow. If necessary, radiation, natural convection, and conduction of heat to mounting surfaces may also be considered. Investigations along these lines are now being conducted by manufacturers and technical committees, resulting either in representative equations or groups of curves, proprietory for individual equipment, which possibly may be standardized later on.

Future Development

It appears that present altitude conditions of 50,000 to 65,000 feet are about the limit for self-ventilation. Beyond this point, the air density decreases to such small values that constant volume fans, designed to pump sufficient air at high altitudes would become prohibitively large at sea level. A fan whose volumetric characteristics would change with altitude would be an acceptable solution, but does not appear to be practical at this moment. What, then, are the possibilities for increased altitude and raised environmental temperatures? The writer believes that the choice among the several solutions must be taken with regard to not only weight, bulk, and efficiency, but also with regard to the logistical point of view, that is, cost, availability, ability to be mass-produced, and reliability and maintenance in the field. The first approach would consist of permitting a reduction of the required life of an inverter to an absolute minimum, both compatible with reliability and over-all cost of flying. For instance, let an inverter for 150 C sea level temperature and 80,000-foot operation be required. Before deciding upon one of the cooling systems to be listed, the writer would give serious consideration to a semistandard inverter, with a reduced life of say, 50 minimum hours at these conditions, which amounts to about 250 to 400 hours' average flying time. The expenditure of a medium-priced inverter after, say, 200 flying hours, is still very reasonable, provided the inverter as such is simple, available in large quantities, and otherwise fairly reliable. Only after this approach is exhausted, more complicated cooling systems should be investigated.

It is fairly safe to say that inverters can be made to operate for a conveniently long period at sea level densities and air temperatures of 120 to 130 C. Consequently, the next approach would consist of an inverter, with a totally enclosed recirculating air system, with the temperatures just mentioned, cooled by an external heat sink, which could be either pressure-controlled water evaporation or fuel-cooling, the latter to be preferred. Should this prove either impossible or insufficient, then water evaporation within the equipment, both stationary and rotating, or fuel-cooling within the equipment should be considered. In stating these methods in such a sequence, the writer by no means wishes to pass judgment on them; however, he feels that the best solution for each condition is the one requiring the simplest design, system- and componentwise.

Having investigated and designed several advanced cooling systems for generators and inverters, the writer wishes to point out that very often the cooling problem, as such, is by no means the unsurmountable problem it may appear at first glance. However, what usually presents much greater obstacles is the difficulty of integrating it into the aircraft system, and there again, this is not so much a technical problem, but rather an administrative one. Co-ordination of information, decision, and implementation among the many engineers participating in a weapon system design, particularly when time is important, has become such a major problem that it very definitely bears upon the technical aspects.

Conclusion

Present-day aircraft inverters appear to have reached their inherent self-cooling capacities with altitudes of 50,000 to 65,000 feet involved. Major difficulties are encountered in the high-temperature characteristics of bearing greases and selenium rectifiers; the brush-wear problem, servicewise, may be solved in an acceptable manner by suitable designs and proper replacement scheduling. For environmental conditions exceeding present ones in severity, several cooling methods are available which should be chosen with regard to inherent characteristics as well as logistical consequences. A fairly acceptable theoretical approach to the airflow and cooling problem is available, which is also suited for rerating questions.

Appendix. Derivation of the Equations

Incompressible Air Flow

FLOW THROUGH ORIFICES

The mass flow through orifice (orifice coefficient-unity) is

$$Q=7.61A\sqrt{\Delta p\times s}\text{ pounds per minute} \qquad (1)$$

For an orderly collection of design and performance data, the following relations prove useful

$$F=\beta\frac{D_0^2\pi}{4}\text{ inch}^2 \qquad (2)$$

$$A=\gamma F\text{ inch}^2 \qquad (3)$$

F is the sum of all duct areas for axial air flow within the machine at the center of the core, and is calculated as a fraction of the total area formed by the outside diameter of the stator lamination. Equations 2 and 3 apply to a-c and d-c generators as well as to duplex flow inverters; for simplex flow inverters

$$A=\gamma\sqrt{\frac{1}{\frac{1}{F_1^2}+\frac{1}{F_2^2}}}\text{ inch}^2 \qquad (4)$$

Typical values for β, γ, and A are:

$\beta=0.09$ to 0.16 up to 0.22, increasing with size

$\gamma=0.40$ to 0.60 depending upon design

A for 2,500 = volt-ampere inverter (simplex) $=2.0\text{–}2.6$ inch^2

AIR FLOW THROUGH A MESH OF DUCTS AND OTHER RESTRICTIONS

The total pressure drop across series-connected restrictions is the sum of the individual drops

$$\Delta p_{tot}=\Sigma\Delta p_i \qquad (5)$$

The individual pressure drop is defined as

$$\Delta p_i=\frac{Q^2}{37{,}300gs}\left(\frac{\xi_i}{F_i}\right)^2\text{ inch H}_2\text{O} \qquad (6)$$

The equivalent orifice of a series connection of restrictions is

$$A_{tot}=144\sqrt{\frac{1}{\sum\left(\frac{\xi_i}{F_i}\right)^2}} \qquad (7)$$

The equivalent orifice of several orifices connected in parallel is

$$A_{tot}=\Sigma A_i\text{ inch}^2 \qquad (8)$$

The value of ξ is defined as the ratio of the particular pressure drop across the restriction to the velocity pressure within the restriction, and thus is a dimensionless constant. For particular values see standard mechanical handbooks.

Pressure drops caused by the forced entrance of air into rotating grids (commutator risers, salient pole wheels) may be approximated by the velocity pressure of the circumferential velocity of such a grid. This drop, by nature, is closer to that of a countervoltage, exerted by a battery in series with a line voltage and several resist-

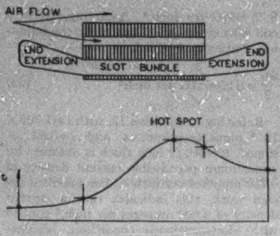


Fig. 1. Temperature distribution

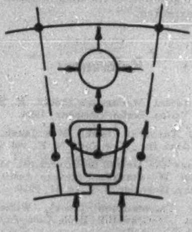


Fig. 2. A-c stator detail

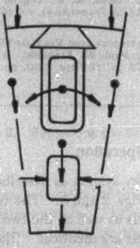


Fig. 3. D-c rotor detail

ances, and not too dependent upon volumetric flow.

Air Mass Flow through a Self-Cooled Inverter

Inasmuch as the cfm are independent of altitude, the mass flow is obtained as

$$Q = \text{cfm} \times s \text{ pounds per minute} \tag{9}$$

The Temperature Field in Aircraft Electric Machines, with Particular Regard to Inverters

It is well known that a stationary or rotating armature cooled by axial air flow may be analyzed mathematically. Assuming circumferential symmetry, differential equations for one slot bundle, within the core and in the end extensions, may be set up. Thus, the temperature distribution may be calculated, as shown in Fig. 1. For an approximate calculation of the gradients resulting in the hot-spot temperature, a simplified calculation may be developed. It is noted that the thermal conductivity across iron laminations is much smaller than along the laminations, with a ratio of roughly 1:50 to 1:60. Because of the small axial temperature gradient in the slot bundle in the vicinity of the hot spot, the heat conducted axially by the copper may be estimated to 10–20%; the remainder must travel radially to the cooling surfaces. A slot segment of unity length, is shown in Figs. 2 and 3. The question of how much heat is either dissipated into or accepted from the air-gap surfaces is, offhand, difficult to answer. However, a comparison of losses on both sides of the gap, as well as available cooling surfaces, will lead to a satisfactory assumption for either an approximate first estimate for a new design or a recalculation of test results. For instance, such a first calculation could be started with the assumption that 25–40% of the rotor losses may cross the air gap. The equations required for such a calculation are self-evident and will not be given here.

Gradients within shunt or exciter field are also predictable by simplified methods. Fig. 4 shows the general picture and the mathematical model, a 2-dimensionally infinite plate with its gradients is shown in Fig. 5. The gradients t_f and t_i may be calculated by conventional methods. In a 2-dimensionally infinite body, with symmetrical heat flow to the two finite surfaces, the internal gradient is

$$t_c = \frac{q}{8k} d^2 \text{ C} \tag{10}$$

For a 1-dimensionally infinite plate, a correction factor is added and we obtain

$$t_c = \eta \frac{q}{8k} d^2 \text{ C} \tag{11}$$

where
$a/d =$
1 ..2 ..3 ..4 ..infinite.circle
$\eta =$
0.59..0.91..0.99..1.0....1.0...0.69

For the specific loss q in watts per inch^3, in copper losses

$$q = \frac{1}{2.54} \epsilon i^2 r_t \text{ watts per inch}^3 \tag{12}$$

and in iron losses

$$q = \frac{2.20 v_{10} s_F}{1{,}000} \left(\frac{B}{64{,}500} \right) m \text{ watts per inch} \tag{13}$$

For the values of the fill factor ϵ and the composite thermal conductivity k the following values are obtained.

For wire wound coil (Fig. 6)

$$\epsilon = \left(\frac{D}{D'} \right)^2 \frac{\pi}{4} \tag{14}$$

$$k = \nu k_I \frac{\text{watts} \times \text{inches}}{\text{C}} \tag{15}$$

where

$D'/D =$
0.5 ..0.6 ..0.7 ..0.8 ..0.9 ..0.95
$\nu =$
1.45..1.75..2.25..3.70..4.75..8.25

For rectangular coil (Fig. 7)

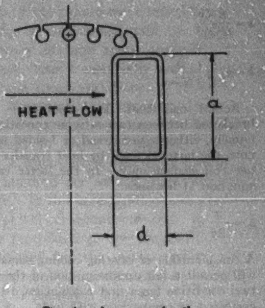


Fig. 4. A-c rotor detail

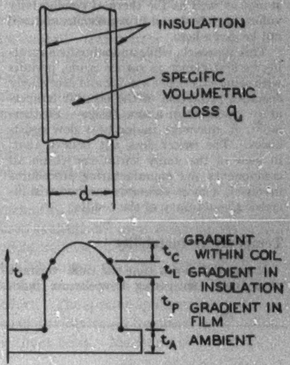

INSULATION
SPECIFIC VOLUMETRIC LOSS q
d
t
t_C GRADIENT WITHIN COIL
t_L GRADIENT IN INSULATION
t_P GRADIENT IN FILM
t_A AMBIENT

Fig. 5. Heat flow in infinite plate

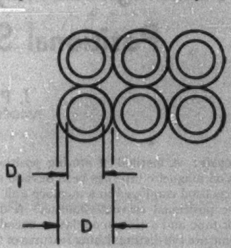


Fig. 6. Round-wire coil

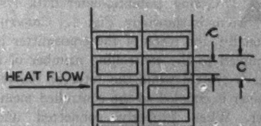


Fig. 7. Rectangular-wire coil

$$\epsilon=\frac{n\times c}{N\times C} \qquad (16)$$

$$k=k_l\frac{c}{C}\,\frac{N}{(N-n)} \qquad (17)$$

Actual coil heat flow conditions are bracketed between two extreme conditions, namely, either symmetrical or 1-sided heat flow, the latter defined by heat transfer by one coil surface only. In the latter case, equation 11 becomes

$$t_c=\eta\frac{q}{2k}\,d^2 \qquad (18)$$

A consideration of existing cooling surfaces will permit a fair approximation of the actual condition for a first calculation, to be corrected by comparison with tests.

Numerical values for heat transfer constants as well as for thermal conductivity values may be found in handbooks and will not be given here.

This approach, while undoubtedly appearing at first glance as rather crude, permits with some experience a quick and acceptably accurate prediction of the possible temperature gradients in a new design. Furthermore, it uncovers major heat flow resistances. The writer does not believe that, in view of the many variations within all components and manufacturing procedures involved, a more elaborate method will increase the accuracy of the results.

Temperature Instability

It may be shown that the time constant of a body, generating temperature independent losses within itself, is equal to

$$T=\frac{Pc_h}{S_a h_a}\text{ seconds} \qquad (19)$$

If, however, the losses within the body are temperature dependent and assumed, in first approximation, proportional to the surface temperature

$$W=W_0[1+\alpha(t-20)]\text{ watts} \qquad (20)$$

then

$$T=\frac{Pc_h}{S_a h_a-W_0\alpha}\text{ seconds} \qquad (21)$$

Thus it is possible that the time constant approaches infinity, that is, the body temperature does not level out. The limiting conditions are described by a maximum value for the specific watt dissipation at cold reference temperature to

$$\left(\frac{W_0}{S_a}\right)\leq\frac{h_a}{\alpha}\text{ maximum permissible watts per inch}^2 \qquad (22)$$

Such a condition may be observed only too often at very high altitudes, with reduced air mass flow and high specific coil losses. A numerical example will illustrate this. Assume a duct 3/16×3/4 inch, an air velocity of 4,000 feet per minute, average air temperature of 80 C, specific air weight of 0.00720 pounds per foot^3 according to an altitude of 50,000 feet; a turbulence factor of 1.8, resulting heat transfer constant of 0.0324 watts per inch,^{2} C and $\alpha=0.00393$, then the maximum permissible specific heat dissipation at room temperature $W_0/S_a=$ 8.25 watts per inch.^{2} Assuming further a coil with symmetrical heat flow, the following relation is obtained.

$$\frac{dq_0}{2}=8.25\text{ watts per inch}^2 \qquad (23)$$

Referring to equation 12, with $r=1.724\times 10^{-6}$ ohms×centimeters, and $\epsilon=0.52$ assumed, a coil 1/2 inch thick is obtained for a maximum permissible current density of 9,700 amperes per inch.^{2} For practical design work, this indicates that a current density of 9,000 amperes per inch,^{2} applied with aforementioned conditions, becomes marginal with regard to temperature stabilization; in addition, a surface temperature gradient of the magnitude of 250 C can be expected.

References

1. ELECTRICAL MACHINERY (book), **R. Richter.** Springer-Verlag, Berlin, Germany, 1924.

2. HEAT TRANSFER (book), **Max Jakob.** John Wiley & Sons, Inc., New York, N. Y., vol. I, 1949.

3. THEORY AND DESIGN OF ELECTRICAL MACHINERY, **Paul W. Franklin.** Stevens Institute of Technology, Hoboken, N. J., pt. II, 1950.

4. HEAT TRANSMISSION (book), **William H. McAdams.** McGraw-Hill Book Company, Inc., New York, N. Y., 1942.

5. SURFACE HEAT-TRANSFER COEFFICIENTS FOR HYDROGEN-COOLED ROTATING ELECTRIC MACHINES, **D. S. Snell, R. H. Norris, Mrs. B. O. Buckland.** *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 174–85.

6. COMMUTATION OF LOW-VOLTAGE D-C AIRCRAFT GENERATORS, **Paul W. Franklin.** *AIEE Transactions*, vol. 72, pt. II, Sept. 1953, pp. 254–62.

A Magnetic Tape Memory For D-C Positional Servomechanisms

J. P. DeBARBER
ASSOCIATE MEMBER AIEE

Synopsis: A method of storing positional data on magnetic tape has been devised and this method employed in a memory unit for a d-c positional servomechanism. A dual track tape and a time modulation encoding scheme are the distinguishing features of this method.

A D-C ring potentiometer data system is shown in Fig. 1. Conventionally, the rotor of the transmitter is coupled to the controlling member of a servo system while the rotor of the receiver is coupled to the controlled member. When properly connected and mechanically installed, the voltage derived from the rotor terminals of the receiver serves as an error signal indicating the magnitude and direction of the misalignment between the controlled member and the controlling member. Ordinarily, this error signal is applied to appropriate amplifiers and drive units to drive the controlled member into alignment with the controlling member. This comprises a basic positioning system.

The relative magnitudes and polarities of any two of the three voltages appearing at the transmitter stator completely specify the position of the transmitter rotor. The third voltage is not independent. Therefore, recording two of these voltages and later playing them into the appropriate terminals of the receiver incorporates an indefinite time delay into the data system. If the controlled member and the controlling member are one and the same, a memory system has evolved.

System Operation

A memory system as described is illustrated in Figs. 2 and 3. The two recording channels as well as the two reproducing channels are identical. The function of the modulator is to produce a square wave of constant repetition rate and to modulate this square wave in such a manner that its average value is proportional to the transmitter voltage being recorded. This is accomplished by a Schmitt circuit triggering on a saw-tooth voltage. The Schmitt circuit is a bistable multivibrator, the transition volt-

---

Paper 54-520, recommended by the AIEE Feedback Control Systems Committee, and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted November 13, 1953; made available for printing October 6, 1954.

J. P. DEBARBER is with the University of Kansas, Lawrence, Kans.

The author wishes to express his gratitude to Prof. Paul C. Cromwell of the University of Tennessee who suggested this project as a Master's Thesis and whose encouragement and advice were major factors in the completion of the investigation. Thanks are also due to Prof. Charles H. Weaver of the University of Tennessee for his interest, unlimited patience, and helpful advice. The author appreciates the efforts of Mr. Williams of the University of Tennessee in building parts for the experimental work.

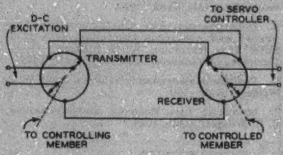


Fig. 1. A d-c ring potentiometer data system

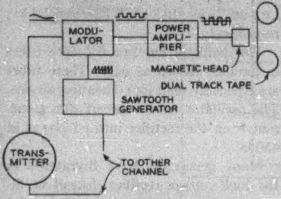


Fig. 2. Recording system

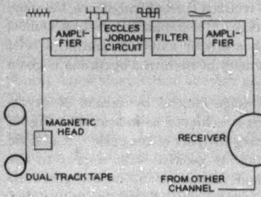


Fig. 3. Reproducing system

age level being a function of the modulating transmitter voltage. The wave-form diagrams of Fig. 4 illustrate the Schmitt circuit operation. The saw-tooth voltage generator is a conventional thyratron relaxation oscillator.

The output of the Schmitt circuit is applied through a power amplifier to the magnetic recording head. The tape is driven to saturation in recording so that provision for erasing is not required since any signal previously recorded is completely obliterated by the new signal.

The process of magnetic tape reproduction is essentially one of differentiation, such that the recorded square-flux pattern gives rise to pulses at the discontinuities. After amplification, these pulses trigger an Eccles-Jordan circuit. The output of the Eccles-Jordan circuit is a square wave, time-modulated exactly as the originally recorded square wave and, consequently, having an average value proportional to the original transmitter voltage. The wave forms associated with the Eccles-Jordan circuit are illustrated in Fig. 5.

The output of the Eccles-Jordan circuit is applied to a clamp circuit which removes the invariant component, amplitude-centering the wave form. This wave form is then acted on by a low-pass filter which passes only the modulating variations which, in turn, are fed to the receiver through a direct coupled amplifier.

The voltage appearing at the receiver rotor terminals is an error signal indicating the magnitude and direction of the misalignment between the receiver rotor and the original position of the transmitter rotor. This error signal, applied to a controller and a motor, drives the receiver rotor into alignment. The drive used in this investigation consisted of a chopper-type amplifier and a 2-phase servomotor.

Since the transmitter and receiver are identical units, one unit can serve in both capacities provided proper electric switching and mechanical coupling are arranged.

Conclusions

A system as described was constructed and it performed quite well. The maximum steady-state positional error encountered was 7 degrees. This error was traced to a noticeably nonlinear sawtooth voltage. The transient performance was not determined quantitatively.

The time modulation method of encoding has several advantages, as shown by Perron.[1] The effect of tape demagnetization is eliminated since the only requirement is that the strength of the recorded flux pattern be sufficient to produce pulses capable of triggering the Eccles-Jordan circuit. Tape demagnetization, being unpredictable in magnitude, practically rules out an amplitude-modulated signal.

Changes in tape speed produce changes in both reproduction level and reproduction frequency and would introduce error into systems employing amplitude or frequency modulation. The time modulation encoding scheme is fairly insensitive to tape speed variations of the type usually encountered. The time-modulated square wave has an average value given by

$$E_{av} = A(2T_0/T-1)$$

where

A = amplitude
T = period
T_0 = time duration of the positive portion of the square wave

This expression is independent of tape speed provided the tape speed is constant over 1 cycle of the square wave. In other words, if the square-wave repetition rate is large in comparison with the highest frequency of tape speed variation, the average value is essentially insensitive to the tape speed variations. The highest tape speed variation frequency likely to be encountered in an acceptable recorder is about 10 cycles per second. Therefore, a square-wave repetition rate of a few hundred per second would be adequate. In this investigation, a repetition rate of 200 per second was used.

In reproduction, the amplitude A is established by the clamp circuit and, consequently, depends only on the stability of the power supplies.

It is conceivable that, with some additional circuitry, this same method could be employed with an a-c synchro generator-control transformer data system. The additional circuitry would take the form of a phase-sensitive detector between the synchro generator and the modulator and a remodulator in the reproducing channel. It is believed that a saturable reactor bridge circuit would serve nicely as a remodulator.

The usefulness of positional servos with memories is self-evident. They could be used as electronic cams and templates and could replace metallic equivalents with the advantage that the pattern could be altered at will with no expense and little effort. The specific applications to which such devices could be fitted are almost without limit. Machines with magnetic

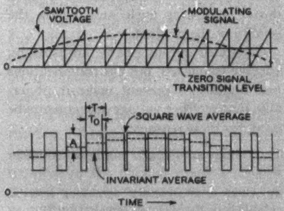


Fig. 4. Schmitt circuit wave forms

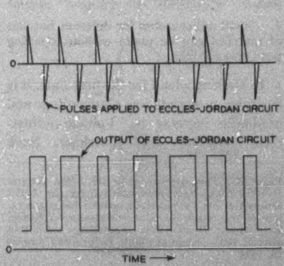

0
PULSES APPLIED TO ECCLES-JORDAN CIRCUIT
OUTPUT OF ECCLES-JORDAN CIRCUIT
0
TIME

Fig. 5. Eccles-Jordan circuit wave forms

memories are likely to play conspicuous roles in industry and national defense.

References

1. Shape Recording With Ratio-Modulated Tape, R. R. Perron. *Electronics*, New York, N.Y., vol. 23, no. 11, Nov. 1950, pp. 104–08.

2. Lead System for Position Memorizing Servomechanism, J. M. Bailey. *Thesis*, University of Tennessee, Knoxville, Tenn., 1951.

3. Magnetic Recording (book), S. J. Begun. Murray Hill Books, Inc., New York, N.Y., 1949.

4. Investigation of a Frequency Memorizing System Controlled by a Direct-Current Voltage, W. O. Campbell. *Thesis*, University of Tennessee, Knoxville, Tenn., 1949.

5. Investigation and Correction of Frequency and Amplitude Instability in a Magnetic Tape Recorder, W. D. Plengey. *Thesis*, University of Tennessee, Knoxville, Tenn., 1949.

6. Position Memorizing Servomechanism, W. P. Walker. *Ibid.*, 1950.

7. Magnetic Tape Program Servomechanism, C. H. Weaver. *Ibid.*, 1948.

Magnetic Amplifier Control of Radio-Frequency Generators

G. R. MOHR
NONMEMBER AIEE

REUBEN LEE
ASSOCIATE MEMBER AIEE

RADIO-frequency (r-f) generators are used in various industrial processes such as soldering, hardening, forging, annealing, etc. The r-f portion of the generator is usually an oscillator comprising one or more large vacuum tubes, a tank circuit, and an inductive coupling or "work" coil. To utilize vacuum-tube capabilities most fully, high voltage is supplied to the tube plates through a rectifier incorporated in the generator. Since high-power vacuum tubes are used, the rectifier is a large one. The control of this rectifier is an important function of the generator, since it affects the power delivered to the load during a heating cycle and, hence, the quality of the product.

Generator output can be adjusted by varying frequency, but the complications and limitations present make it preferable that rectifier voltage adjustments be used, for these reasons:

1. Voltage adjustment can be made under load.
2. Once a satisfactory operating level is established, it can be duplicated easily.
3. If necessary, wide range control can be had.
4. Small adjustments for different batches of material can be made without opening generator cabinet.

In starting a load for the first time, it is desirable to work at low plate voltage and gradually increase it. Voltage adjustment simplifies load matching. Such voltage adjustment enables the operator to study heat patterns or to work toward a maximum loading of the generator. Very often a number of different jobs can be done using the same load coil simply by changing the power level. Voltage adjustment, therefore, increases the ease and versatility of operation. Simplicity of operation, ease of maintenance, and adaptability to a large variety of applications are desirable attributes of industrial heating equipment.

Approximate voltage adjustments are made, in steps, by taps on the plate transformer. Transformer taps have a low first cost, but have enough disadvantages so they are rarely used for a complete system of voltage control. The main disadvantages are:

1. Some means of changing taps must be provided.
2. Doors on the rectifier have to be unlocked to change taps. Where one man is responsible for keys, this is inconvenient.
3. Taps cannot be changed under load.
4. Voltage can be adjusted in steps only and, therefore, it may not be practicable to load the generator for maximum plate current at all times. This also is inconvenient when setting up new jobs.

For smoother regulation of plate voltage, various control systems are used.

The generator of Fig. 1 is designed to take power from a 230 or 460-volt 60-cycle line and deliver 25-kw r-f power into the work load at a frequency of 450 kc. A circuit diagram for the generator is shown in Fig. 2. D-c power for the oscillator is supplied by a 3-phase full-wave rectifier employing six *WL-575A* diode tubes, V_1 to V_6. Plate transformer secondary windings T_1, T_2, T_3 have taps so that four different power levels can be obtained. Plate voltages at the four taps are 12 kv, 10.2 kv, 8.7 kv and 7.4 kv, respectively.

The oscillator has two *WL-5668* tubes in parallel, shown as one. These tubes are especially built for r-f heating service. The oscillator tubes convert direct current from the rectifier into r-f power at 450 kc.

Many generators have thyratrons in the high-voltage rectifier instead of diodes. A phase-shifting network is used in the thyratron grid circuit to control closely the phase back or firing angle of the rectified current and, hence, the plate voltage. To compare thyratron control with magnetic amplifier control, a brief summary of thyratron operation is given here.

Voltage control by means of thyratrons is achieved as indicated in Fig. 3. During the part of the cycle when anode voltage is positive with respect to the cathode, anode current flows if grid voltage is more positive than the critical grid voltage. Once thyratron conduction is started, change of grid voltage to a value lower than critical does not stop conduction. Conduction does stop, however, when the anode voltage falls to a value less than that of the cathode. Cathode

Fig. 1. 25-kw r-f generator

Paper 54-377, recommended by the AIEE Electric Heating Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted May 5, 1954; made available for printing August 16, 1954.

G. R. Mohr and Reuben Lee are with the Westinghouse Electric Corporation, Baltimore, Md.

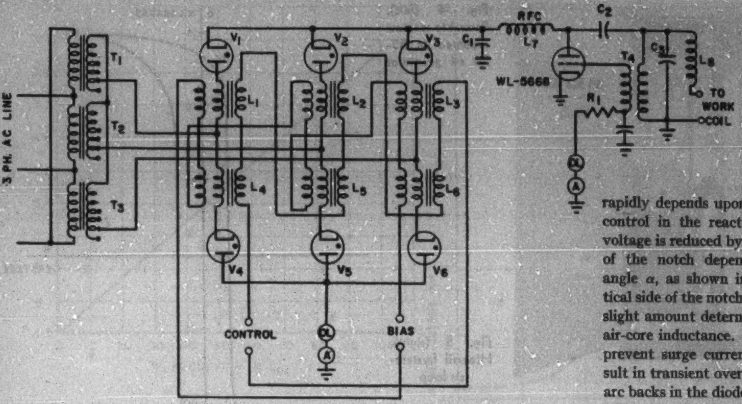


Fig. 2. Schematic diagram of 25-kw r-f generator

voltage is shown as zero in Fig. 3. Thyratron current may thus be made to start once each cycle at the instant when total grid voltage exceeds critical value.

Grid voltage has a negative d-c component, lower than critical at every point. The phase position of a superposed alternating voltage is changed by a phase-shifting network. Increasing the phase angle of the a-c grid voltage delays the angle at which conduction begins and, hence, reduces the thyratron current from maximum. The shaded area indicates the portion of the cycle during which thyratron current flows. At the instant of firing, current rises abruptly from zero to a value dependent on the instantaneous anode voltage. This sudden rise of current may cause a surge of voltage, and damage the rectifier insulation. Surge protective means must then be employed. Moreover, in high-voltage rectifiers, if thyratron arc backs occur, the grid circuit is subjected to high voltage so that the grid control circuits must all be insulated for this voltage.

Magnetic Amplifier Operation

Plate voltage control is greatly simplified when saturable reactors are used. In the r-f generator described, six reactors control approximately 50 kw of rectified power and require 24 watts of control power. These reactors and the associated rectifiers constitute a self-saturated magnetic amplifier with power gain of more than 1,000 to 1. The reactors are shown as L_1 to L_6 in Fig. 2, connected in series with the high-voltage diodes. These diodes are needed for plate voltage rectification in any case. Placing the control reactors in series with the diodes causes the rectified current to saturate the reactors. This greatly increases the effectiveness of the magnetic amplifier, as is amply described in references 1, 2, and 3. The reactors must be insulated for the full voltage of the rectifier. If they were connected in the primary circuit, the insulation could be reduced, but additional rectifiers or much larger reactors would be necessary for the same degree of control.

The six saturable reactors, L_1 to L_6 of Fig. 2, are mounted in the base of the r-f generator, as shown in Fig. 4. These reactors weigh a total of 252 pounds, which is 30 per cent of the weight of the rectifier plate transformer. The reactor cores are Hipersil grain-oriented silicon steel which has found wide use in transformers of all kinds. A property of Hipersil which is particularly valuable in saturable reactors is the shape of its hysteresis loop; see Fig. 5. The more nearly rectangular and narrow the hysteresis loop, the more efficiently the core material functions in a saturable reactor. With a narrow loop, the control current and power required to change the plate voltage are smaller.

The steep part of the hysteresis loop results in high reactor inductance which effectively delays the build-up of current through the rectifier tubes in a manner similar to the firing-angle delay of the thyratron. When the reactor flux density builds up to the value B_s in Fig. 5, the reactor inductance suddenly becomes small, limited mostly by the reactor air-core inductance. Then, there is a rapid increase of current. The phase angle at which the current begins to increase rapidly depends upon the amount of d-c control in the reactor. Rectified plate voltage is reduced by phase back because of the notch dependent on the firing angle α, as shown in Fig. 6. The vertical side of the notch actually slopes by a slight amount determined by the reactor air-core inductance. This slope helps to prevent surge currents which might result in transient overvoltages and, hence, arc backs in the diodes.

It may be observed from Fig. 6 that the maximum value α can have is 120 degrees. At 60 degrees, the next reactor begins absorbing voltage but the first reactor may continue to absorb it also. If all six reactors were biased for a firing angle of 120 degrees, there would be nearly zero voltage output from the rectifier. For firing angles intermediate between zero and 120 degrees, the voltage is reduced from maximum voltage E_m, according to the upper curve of Fig. 6. If it is desired that the reactor absorb a certain fraction of rectified output voltage, the reactor size may be found from Fig. 7. This is the relation between turns in a saturable reactor and turns in a plate transformer with the same core, voltage, flux density, and frequency. Fig. 7 therefore gives reactor size in terms of plate transformer size. It is the plot of equation 5 derived in the Appendix.

Each reactor has a bias and a control winding. These windings have the same number of turns. The ampere-turns required in the control winding for full output would be approximately the same as in the load winding if simple saturable reactors were connected in the transformer primary or secondary circuit.

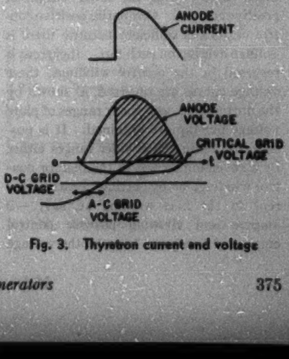


Fig. 3. Thyratron current and voltage

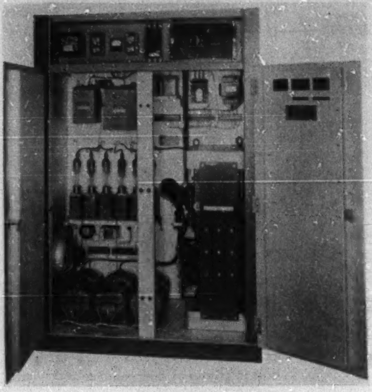


Fig. 4 (left). Saturable reactors mounted in 25-kw r-f generator

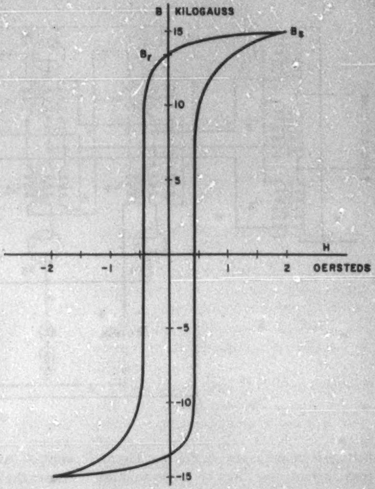


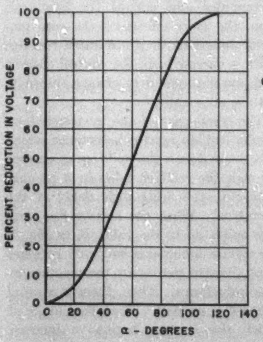
Fig. 5 (right). Hipersil hysteresis loop

Connecting the reactors in series with the rectifiers, as shown in Fig. 2, reduces the control ampere-turns to approximately 17 per cent of the load winding ampere-turns.[1,2] The load winding is insulated for full rectifier output voltage to ground and to the control windings. Bulky control winding insulation is avoided by using a small number of turns and control current of the same order as the rectifier output current.

Control windings of all six reactors are connected in series, and so are the bias windings. Usually the bias windings are short-circuited by placing a jumper across the bias terminals, as shown in Fig. 2. The short circuit reduces ripple harmonics in the rectified output of the control d-c supply. The short-circuited bias winding thus acts as a ripple filter. It also increases the range of voltage control. Changes in power level are controlled by inserting a low-voltage d-c source and potentiometer in the windings connected to the control terminals. Typically smooth voltage control curves are shown in Fig. 8. For each plate transformer tap, voltage rises at the same rate with increasing positive control current, more rapidly with the lower values of positive current. Ordinarily positive control current is sufficient because there is voltage overlap on each tap. If current is reversed in the control windings, these voltage curves are reversed, as shown by the dotted lines, and wider ranges of plate voltage per tap are obtained. It is possible to obtain these wider ranges either by a reversing switch in the reactor control windings or by setting negative current in the bias winding at, say, −2 ampere and allowing positive control current of 4 amperes to cause the voltage to rise. If the range of plate voltage is restricted to the region of about zero control current, shown in Fig. 8, approximately linear control may be had. Approximate generator outputs in kilowatts for these same control currents and voltage ranges are shown in Fig. 9.

A further distinction between the curves of Fig. 8 and those of Fig. 9 is the fact that Fig. 8 is a plot of voltage across a constant impedance. Maximum output power on tap no. 4 is 11,600 volts times 4 amperes direct current or 46.4 kw. Minimum power on this tap is 7,400 volts times (4×7,400/11,700=2.55 amperes) or 18.9 kw. Rectifier output power range is 46.4−18.9=27.5 kw. Generator or heating load power is less than the rectifier power because of oscillator losses. This is illustrated by Fig. 9, in which the range of generator output power differs considerably from that of Fig. 8.

No-load to full-load regulation curves with fixed direct current in the control winding are shown in Fig. 10. With 3 amperes positive d-c control current, the curve has the usual shape of a rectifier regulation curve with slightly drooping voltage regulation. Reducing the control current reduces the plate voltage and causes a greater amount of regulation from zero to 4 amperes d-c plate current. However, from 1- to 4-ampere plate current, the range of practical interest because the oscillator no-load plate current is usually about 1 ampere, the regulation curves have about the same slope for all values of control current.



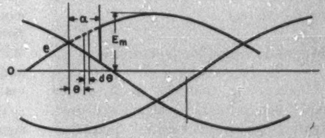


Fig. 6. Rectifier plate voltage versus firing angle

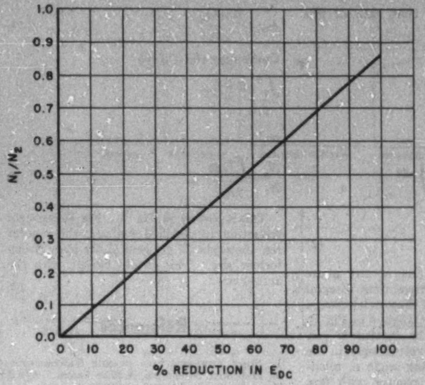


Fig. 7 (above). Reactor size versus reduction in rectified voltage

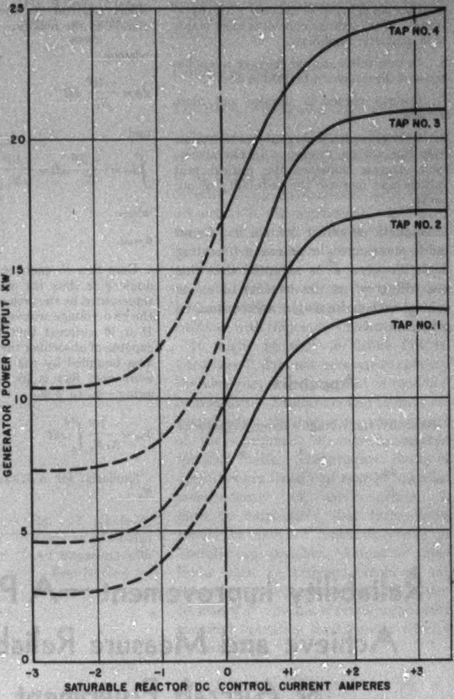


Fig. 9 (right). Power output versus control current

Response time, or the time for a magnetic amplifier to assume steady-state conditions after a change is initiated, plagues many magnetic amplifier applications. Conditions in r-f heating apparatus are more favorable. Response time should not be so long as to interfere with short-time heating cycles. Response time for the 25-kw generator is 0.35 second; it can be made shorter, in unusual instances, by the addition of series resistance in the control circuit. This requires more control power; it is a disadvantage of magnetic control, but it is negligible in most heating applications.

The advantages of magnetic amplifier control compared to thyratron control are the following:

1. Slightly sloping wave front at the firing instant provides a cushion against transient voltages. This eliminates the need for surge-suppression networks used with thyratron control. Efficiency is better because of surge-suppression resistors and associated losses are eliminated.

2. Because of this sloping wave front, as each diode fires, the ratio of peak-to-average voltage is lower, and the work load may be coupled tighter without causing flashovers between the work piece and the work coil.

3. Diode tube life is much longer than thyratron life. This is further enhanced by the cushion mentioned in item 1 and simplifies maintenance materially.

4. Diode tubes are materially lower in cost as compared to thyratrons.

5. Firing angle is not influenced by strong

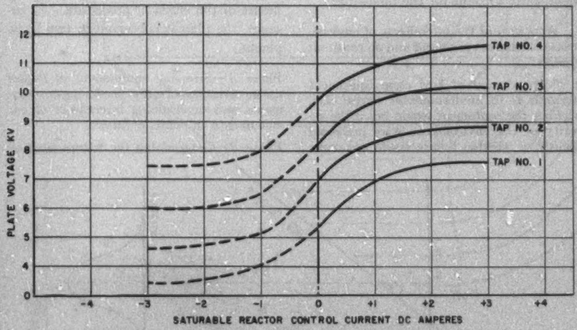


Fig. 8. Plate voltage control with combined bias and plate current

- - - - Negative control current
—— Positive control current

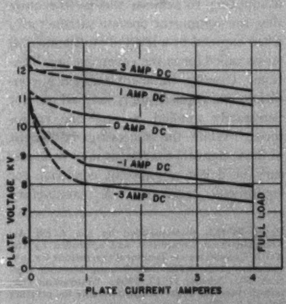


Fig. 10. Regulation curves in rectifier with saturable reactor control

neighboring fields as is the case with thyratrons. Thus, voltage control is more stable with magnetic amplifiers.

6. In case tubes arc, the damage is less because of the cushion described in item 1.

7. Control circuit is simpler and more flexible.

8. Voltage control is less subject to trouble from line voltage changes. In thyratrons, such changes influence the critical grid voltage and increase the possibility of arc backs.

Magnetic amplifier control has found quick acceptance in several r-f heating fields already. It is expected that this use will grow as the benefits of easier control and maintenance are realized in practice.

Appendix

Basically, the voltage e across a reactor of N_1 turns is

$$e = -N_1 \frac{d\phi}{dt} \times 10^{-8} \tag{1}$$

where $(d\phi)/(dt)$ is the time rate of flux change in the reactor.

whence

$$d\phi = \frac{-10^8}{N_1}\, e dt$$

and

$$\int d\phi = \frac{-10^8}{N_1} e dt = \frac{-10^8}{\omega N_1}\int e d\theta$$

where

$$\theta = \omega t \tag{2}$$

Core flux in each reactor at any phase angle θ is thus the voltage time integral represented by the area of the notch between the two voltage waves at angle θ in Fig. 6. If it is assumed that the reactor is just capable of absorbing the volt-seconds in the area bounded by the firing angle α, maximum core flux is for a reactor with flux swing $-\Phi_m$ to $+\Phi_m$

$$\Phi_m = \frac{-10^8}{2\omega N_1}\int_0^\alpha e d\theta \tag{3}$$

Similarly, for a transformer with turns N_2

$$\Phi_m = \frac{-10^8 E_m}{\omega N_2} \tag{4}$$

Combining these gives

$$\frac{\int_0^\alpha e d\theta}{2N_1} = \frac{E_m}{N_2}$$

or

$$\frac{N_1}{N_2} = \frac{\int_0^\alpha e d\theta}{2E_m} \tag{5}$$

This is plotted in Fig. 7. For 36-per-cent reduction in rectified voltage, the reactor size should be 31 per cent of the plate transformer size. This checks closely with the actual size.

References

1. Some Fundamentals of the Transductor or Magnetic Amplifier, A. Uno Lamm. *AIEE Transactions*, vol. 66, 1947, pp. 1078–85.

2. Self-Saturation in Magnetic Amplifiers, W. J. Dornhoefer. *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 835–50.

3. Steady-State Analysis of Self-Saturating Magnetic Amplifiers Based on Linear Approximations of the Magnetization Curve, Walter H. Esselman. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 451–59.

Reliability Improvement—A Plan to Achieve and Measure Reliability of Aircraft Equipment

J. E. LUCKMAN
ASSOCIATE MEMBER AIEE

All procurements and procurement specifications are made with an attempt to obtain equipments that are reliable. The present, classical method of attempting to achieve this is to require that the equipment operate satisfactorily when exposed to a level of environmental stress called for in the governing specification. Usually, these environmental conditions are applied singly, and all are not necessarily applied to any one equipment. While the successful fulfillment of these classical conditions may be necessary, it is not a sufficient condition for equipment reliability. A satisfactory plan for the development of reliable equipment must account for the following:

1. Appraisal of the intereffects of environments on stresses produced and on resultant failures.

2. Measurement of how close equipment strength is to environmental stress; i.e., whether the equipment would fail if the severity of the environment were increased slightly or whether failure would not occur until the severity of the environment were increased many times.

3. Normal variations in the strength of the equipment; i.e., there must be assurance that it was not a random chance that a particular equipment successfully withstood a specific environmental stress.

4. Appraisal of the effects of normal variations of the magnitude of environmental stresses in specific applications.

The New Plan

The proposed method of achieving and measuring reliability rectifies the weaknesses of the classical procedure. In essence, the plan is conceived in two basic phases.

Phase *A*—Stressing equipments to failure under variations of stress-producing environments and determining intereffects of environments on points of failure.

Phase *B*—Determining the nature such as

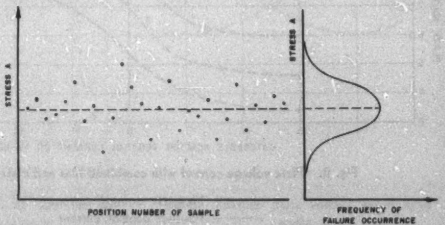


Fig. 1. Distribution of failure with respect to stress

Paper 54-353, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 10, 1954; made available for printing August 2, 1954.

J. E. Luckman is with the United States Naval Air Development Center, Johnsville, Pa.

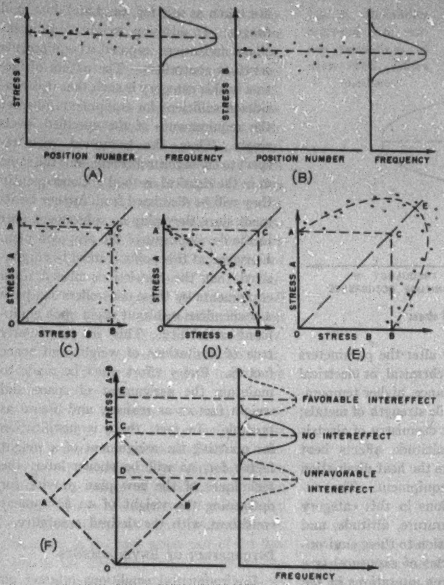


Fig. 2 (left). Intereffect of environments

Fig. 3 (below). Variation of median and dispersion with time (cycles)

magnitude, frequency, and distribution of environments for any particular application.

Details of Phase A

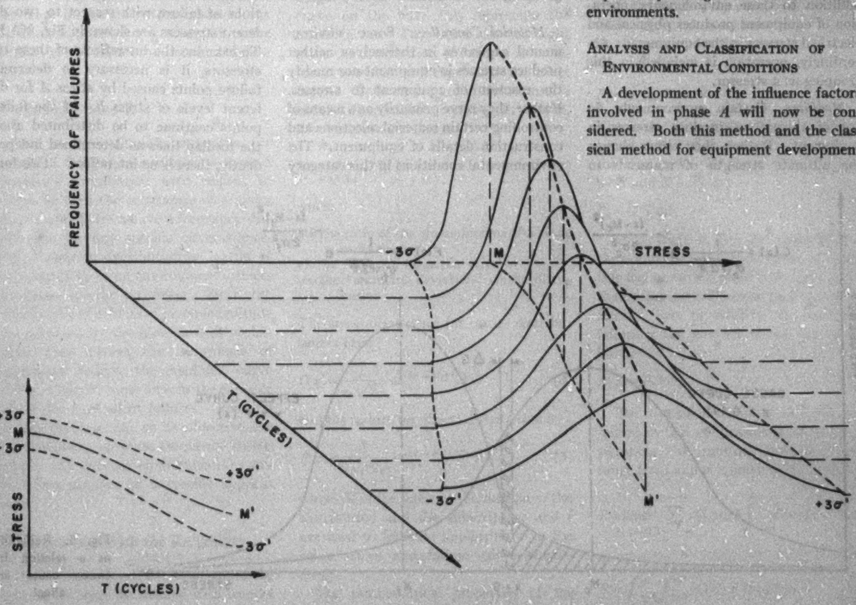
The Basic Stress-to-Failure Principle

The new plan departs from the classical method in that the equipment is subjected to an increasing stress by increasing the severity of the environment until a failure occurs in the equipment. In Fig. 1 the plot on the left shows the magnitude of the stress at which failure occurred for a number of units. This information is projected and shown on the right as a frequency distribution of the failure occurrence with respect to stress magnitude.

It might be well to define failure. Used here it does not necessarily connote the idea that the equipment is rendered inoperative. A failure is said to have occurred when the operating performance of the equipment is outside required specified limits. The required limits of operation are those that may be specified independently of environments. It must be emphasized that these limits should be assigned realistically and as liberally as possible. Actual required limits must be assigned, avoiding any tendency to assign narrower limits of performance as a means of retaining actually required operating limits under adverse environments.

Analysis and Classification of Environmental Conditions

A development of the influence factors involved in phase *A* will now be considered. Both this method and the classical method for equipment development



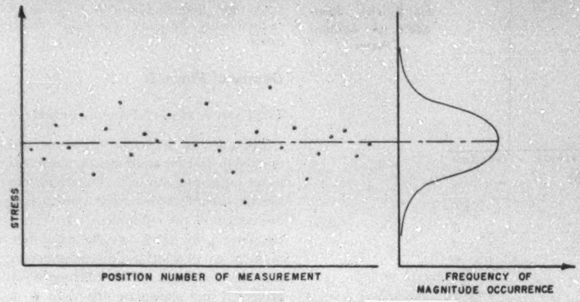


Fig. 4. Distribution of magnitude of an environmental stress

draw the influencing environmental conditions from the same circumstances of aircraft application. These influencing environments are taken as vibration, shock, acceleration, temperature, altitude, humidity, salt fog, fungus, and sand and dust. For the purpose of this paper, these environmental conditions are divided into three categories nominally termed stressors, modifiers, and material controllers.

Stressors. Conditions which in themselves produce stresses in equipments are classified as stressors. The environmental conditions in this category are taken as vibration, shock, and acceleration. In addition to these environments, operation of equipment produces physical and electrical stresses in the equipment. Accordingly, operation is included in this grouping as a stressor.

Modifiers. Certain environments do not of themselves produce stresses in equipment. Rather, they either change the ultimate strength of materials to resist stress or they alter the parameters of certain physical, chemical, or electrical reactions. For instance, higher temperatures lower the tensile strength of metals; moisture affects the chemistry of electric brush operation; altitude affects heat transfer and so alters the heat dissipation characteristics of equipment. The environmental conditions in this category are taken as temperature, altitude, and humidity. In addition to these environments, time (or cycles of exposure) is a factor which produces alterations in the ability of equipment to resist stresses. Accordingly, time is included in this grouping as a modifier.

Material Controllers. Some environmental exposures in themselves neither produce stresses in equipment nor modify the reaction of equipment to stresses. Rather, they serve primarily as a means of controlling certain material selections and construction details of equipment. The environmental conditions in this category are taken as salt fog, sand and dust, and fungus. In addition, weight and dimensional limitations imposed on equipments serve as controllers. The nature of factors in this category is such that it is considered sufficient for equipments to meet the requirements of the specified levels singly. Thus, the treatment with respect to these controllers remains the same as in the classical method. Consequently they will be dismissed from further treatment since they play no subsequent part in the development of the proposed plan. However, at this point it must be emphasized that the restrictions placed upon equipments by these controllers can have a tremendous resultant effect upon equipment reliability. This is particularly true of limitations of weight and space factors. Every effort must be made to maintain the assignment of space and weight factors as realistic and liberal as possible. In fact, there is justification for omitting the assignment of a weight factor for, as will be shown later, the techniques of the new plan provide for optimizing the weight of an equipment consistent with the desired reliability.

INTEREFFECT OF ENVIRONMENTS

Environmental conditions interact on equipment performance, and such intereffect on equipment failure must be investigated. For example, the distributions of failure with respect to two different stressors are shown in Fig. 2(A,B). To examine the intereffects of these two stressors, it is necessary to determine failure points caused by stress *A* for different levels of stress *B*. If the failure points continue to be distributed about the median lines as determined independently, there is no intereffect. This form

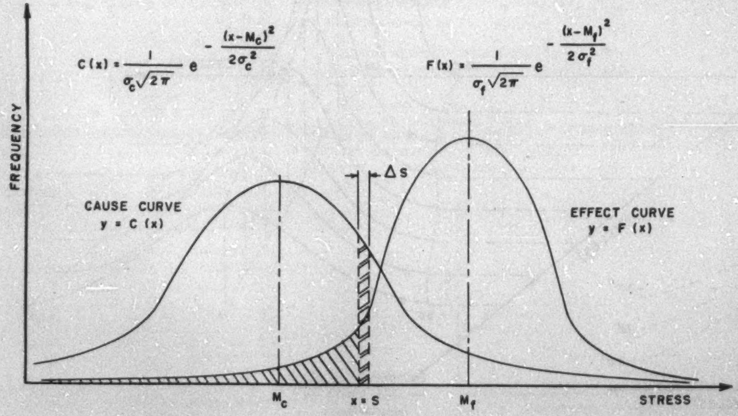


Fig. 5. Reliability as a relation between cause and effect

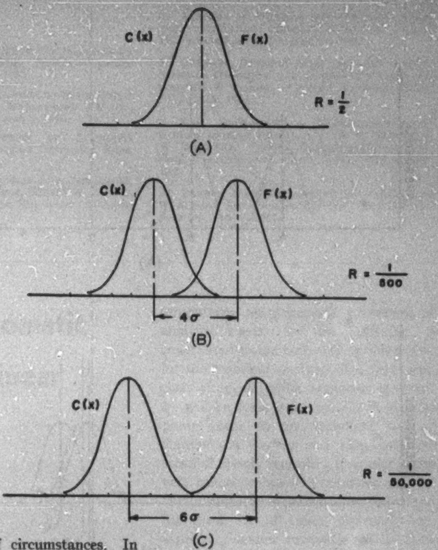


Fig. 6. Reliability achieved with various separations of stress and failure medians

of distribution is shown in Fig. 2(C). However, if the distribution trend is below these medians, as is illustrated in Fig. 2(D), there is an unfavorable intereffect of stressors *A* and *B*. On the other hand, if the trend is above the previous medians, as shown in (E), then there is a favorable intereffect. An expedient way of determining such intereffects is to vary both stressors simultaneously. Let the line *OC*, shown in Fig. 2(C, D, E) define a new stressor *A-B*. Graph (F) shows the application of the basic stress-to-failure principle to determine the intereffects of *A* and *B*. If the median of the distribution of failures caused by stressor *A-B* falls at *C*, there is no intereffect between *A* and *B*. If the level is lowered to *D*, there is an unfavorable or adverse intereffect, and if the level is raised to *E*, there is a favorable intereffect.

While two stressors were used here to illustrate the point, this principle is the same if a stressor and a modifier are combined. Neither is the principle limited to the combination of two factors. Two elements only are used here for simplicity of illustration. Of particular interest, however, is the case where one of the factors is the modifier, time (or cycles), an ever-increasing quantity. Fig. 3 shows how the median and distribution of failure points may be expected to vary with time. The graph at the lower left shows the trace in the time-stress plane of the median and points 3 σ on either side of the median of the failure frequency distribution.

Details of Phase B

Just as the ultimate strengths (or points of failure) of equipments have a frequency distribution with respect to stress, so does the occurrence of a given magnitude of stress have a frequency distribution for any specific circumstance. Fig. 4 shows a distribution or scatter of the magnitude of an environmental stress for some specific condition. Since the occurrences of a stress of given magnitude are governed by the same laws of probability that govern the occurrence of equipment failure, the resultant curve of Fig. 4 has the same form as the curve of Fig. 1 for equipment failures. Phase *B* of this program has as its objective the determination of stress frequency distribution and definite environmental combinations peculiar to a specific application.

A Measuring Stick for Reliability

Reliability is a relative term and to be meaningful must be defined with respect to a specific set of circumstances. In terms of the type of data discussed in the *A* and *B* phases of this program, consider the relations existing between cause and effect. In Fig. 5, the distribution of magnitude of stresses is shown as the curve on the left; this represents the cause factor. The curve on the right gives the distribution of failures with respect to stress and represents the effect factor. Both curves are in the form of a normal distribution or probability curve. The general equation is

$$y = \frac{1}{\sigma\sqrt{2\pi}} \epsilon^{-x^2/2\sigma^2} \qquad (1)$$

where

π = the ratio of the circumference of a circle to its diameter
ϵ = the Naperian logarithm base
σ = the "standard deviation" of the distribution

The stress occurrence (or cause) distribution is then

$$C(x) = \frac{1}{\sigma_c\sqrt{2\pi}} \epsilon^{-(x-M_c)^2/2\sigma_c^2} \qquad (2)$$

and the failure (or effect) distribution is

$$F(x) = \frac{1}{\sigma_f\sqrt{2\pi}} \epsilon^{-(x-M_f)^2/2\sigma_f^2} \qquad (3)$$

where *M* is the value of the median of the distribution and the subscripts *c* and *f* are used to indicate applicability to the cause curve and failure curve respectively.

The mathematical properties of the probability curve are such that if any point *x* (designated on Fig. 5 by the vertical line *S*) is taken, then the probability that the applicable equipment will fail when subjected to such a stress *S* is given by the area under the failure curve up to the value of $x = S$. This is expressed by

$$P_f = \int_{-\infty}^{S} F(x)\,dx \qquad (4)$$

The probability that a stress of magnitude between *S* and $S + \Delta S$ will occur is the area under the cause curve bounded by *S* and $S + \Delta S$ or

$$P_c = C(S)\Delta S \qquad (5)$$

Thus, there is first the probability of the occurrence of a given stress, and after this has occurred, there is the probability that a failure will occur because of this stress. The probability of equipment failure is the product of the two probabilities

$$R(S) = P_c P_f \qquad (6)$$

The over-all probability of equipment failure in this application is obtained by summing this product over the entire range from minus infinity to plus infinity

$$R = \underset{\Delta S \to 0}{\text{limit}} \sum_{S=-\infty}^{S=\infty} C(S)\Delta S \int_{-\infty}^{S} F(x)\,dx \qquad (7)$$

or

$$R = \int_{-\infty}^{\infty} \int_{-\infty}^{S} C(S)\,F(x)\,dS\,dx \qquad (8)$$

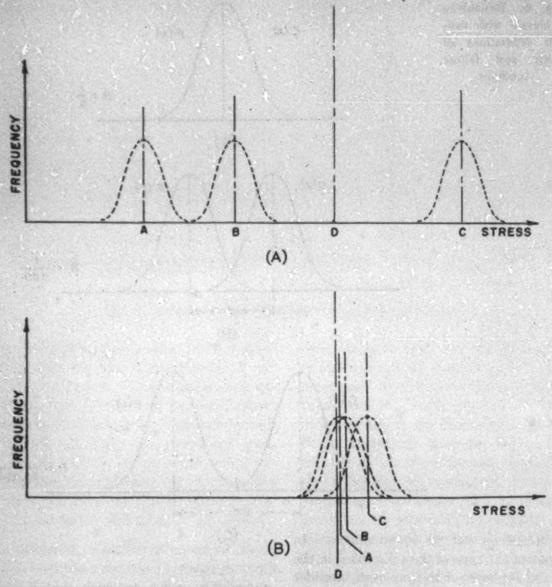


Fig. 7. Effect of reliability concept on equipment improvement

This is a measure of the reliability of the equipment in this application. The double integral may be reduced to the form

$$R=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_e^2+\sigma_f^2}}\times \int_0^{\infty} e^{-\frac{[x+(M_f-M_e)]^2}{2(\sigma_e^2+\sigma_f^2)}}\,dx \quad (9)$$

which may be evaluated from standard probability tables.

This, then, is the definition of reliability as conceived in this program. Numerically, it represents the number of failures, 1 per 100 or 1 per 1,000, that may be expected for the specific application. It will be recognized that this is of the nature of a calculated risk. It will be seen that reliability is dependent upon the dispersion of the two sets of data and the separation of the medians.

Certain factors are very evident if the reliability of equipment is examined under this concept. Fig. 6 shows several examples. In Fig. 6(A) is illustrated a case where the limits of stress specified coincide with the median of the stresses, and the median of equipment failure coincides with the same value. Under the classical method, equipment so designed has a 50–50 chance of being qualified as satisfactory. In actual service use one failure out of every two equipments could be expected. This reliability of 1 per 2 could be predicted by the proposed measurement method. The curves in Fig. 6(B) represent a condition where the medians are separated by four standard deviations. Reliability in this case is approximately 1 per 500. In (C), the separation is six standard deviations and the corresponding reliability is 1 per 50,000. For a separation of seven standard deviations, reliability is less than one failure per 3,000,000 equipments. These illustrations are used to demonstrate the fact that the proposed plan is capable of detecting not only areas of underdesign but also areas of overdesign.

Implementation of the Program

It is evident that phase *A* of this program can be initiated independently of phase *B*. Phase *A* is a laboratory program that should be carried out concurrently with the development of an equipment. Phase *A* can be a powerful development tool when applied in this manner. Figure 7(A) illustrates in part this scheme of application. The line *D* represents the desired median of strength for the equipment. Curves *A*, *B*, and *C* represent the failure distributions obtained where the element causing failure of the equipment was component *A*, *B*, and *C* respectively. The subsequent curves of failure at increased stress were obtained either by strengthening the weak component or removing it from the effects of the environment and continuing the investigation at higher stresses. Figure 7(B) shows the results obtained after further development of the equipment strengthened components *A* and *B* and brought component *C* more in agreement with the over-all requirement. By this type of procedure, factors such as weight of the equipment can be optimized.

It is also evident that phase *B* is a necessary part of this program if it is to achieve the greatest utility of application. However, phase *B* is a field program rather than a laboratory program. Flight service and aircraft manufacturers could contribute to the formulation of data for this phase. Meanwhile, the execution of phase *A* could be pursued and its results applied under the assumption that presently specified limits represent the medians of actual stress distributions. This assumption would be subject to later correlation with data from phase *B* and with data returned from the fleet on actual service success.

Conclusions

In the execution of this plan, the program must be resolved into individual projects. Details within any specific project will be guided by the over-all concept of the general approach and will be subject to statistical analysis and probability studies to assure effective coverage of the problem. When applied to new developments or equipment improvements, for military applications, the reliability factor *R*, which represents a calculated risk, should not be left to the discretion of individual engineers or designers, but rather preassigned uniformly for a given group of equipments by top military authority.

This plan for the achievement of equipment reliability approaches the over-all problem in the true systems concept. Resultant equipment improvements resulting from the application of this program will go a long way to reduce maintenance problems on equipments because it will automatically drive toward maintenance-free units. When the reliability measurement scheme evolved is applied to information obtained from phase *A* and with respect to the specific reference levels determined by phase *B*, the system becomes an important and powerful tool. It will be capable of the following:

1. Determining areas and specific equipments where further development is necessary.

2. Assisting in the development and improvement of specific equipments by segregating and emphasizing specific details of underdesign and overdesign.

3. Giving an advance measure of dependability for any specific equipment in a particular application.

References

1. Some Uses of Statistics in the Planning of Experiments, C. A. Bicking. *Industrial Quality Control*, vol. X, no. 4, Jan. 1954.

2. Concepts and Applications of Reliability, R. R. Carhart. *Research Memorandum RM-531*, The Rand Corporation, March 1, 1951.

3. Mathematical Methods of Statistics (book), H. Cramer. Princeton University Press, Princeton, N. J., 1946.

4. Techniques of Statistical Analysis (book), edited by Churchill Eisenhart, Millard W. Hastay, W. Allen Wallis. McGraw-Hill Book Company, New York, N.Y., 1947.

5. A Study of Methods for Achieving Reliability of Guided Missiles, R. Lusser. *NAMTC Technical Report 75*, July 19, 1950.

6. General Specifications for the Safety Margins Required for Guided Missile Components, R. Lusser. *NAMTC Technical Report 84*, July 10, 1951.

7. The Experimental Determination of Guided Missile Reliability, M. R. Seldon, D. W. Pertschuk. *Journal*, Operations Research Society of America, vol. II, no. 1, Feb. 1954, pp. 31–40.

8. Project Rand Reliability Reference List A. *Research Memorandum RM-529*, The Rand Corporation, Feb. 1, 1951.

Phase-Plane Analysis of Automatic Control Systems with Nonlinear Gain Elements

RUDOLF E. KALMAN
STUDENT MEMBER AIEE

Synopsis: This paper develops a unified method of transient analysis of second-order feedback control systems having nonlinear gain components and subjected to step or ramp inputs. Essentially, the technique is a decomposition of the phase plane into linear regions within which trajectories converge to corresponding critical points. Application of the method to several typical problems previously analyzed by other means shows that considerably more insight may be obtained.

THIS paper presents in a new light some of the basic concepts of nonlinear mechanics relating to the analysis of feedback control systems; the objective is to develop a simple but penetrating qualitative theory for the transient behavior of such systems.

In recent years, especially since the availability of Minorsky's excellent monograph,[1] there has been considerable interest in the application of the so-called phase-plane method to the solution of otherwise unwieldy problems encountered in connection with servomechanisms containing a relay,[2,3] backlash,[4,5] or Coulomb friction.[6,7]

The most important advantage of the phase-plane method is that it displays all possible solutions of a given differential equation. To appreciate this point, it should be recalled that any linear feedback system may be uniquely described in terms of a single input-output relationship; e.g., the response to a step input signal. This is, of course, equivalent to saying that such a system may be described by a transfer function in the frequency domain. No more information is necessary to determine the response of the system to any input signal, although details of the calculation may be tedious. All of the foregoing properties of a linear system may be shown to follow directly from the principle of superposition which, by definition, is not applicable to a nonlinear system. Instead of a single input-output pair, a large number of pairs (strictly speaking, infinitely many) are needed to characterize a nonlinear system. As the amplitude of the input signal increases, there may be essential changes in system behavior; it is, therefore, desirable to have at least a qualitative picture of all possible responses of a nonlinear system for certain types and amplitude ranges of input signals. The phase-plane representation is the natural medium for the investigation of such problems.

The disadvantage of the phase-plane method is its restriction to second-order systems on the one hand, and step or ramp inputs (or combinations of these) on the other. These limitations are not inherent in the fundamental concept but merely stem from the difficulty of dealing with the geometry of multidimensional spaces.

Present-day literature on the applications of the phase-plane concept[2–7] suffers greatly from the lack of unifying principles. Problems are being attacked separately; no systematic procedure of analysis exists. In the following, a generalized framework will be developed for the transient analysis of a very wide class of second-order nonlinear systems. As will be seen, it is possible to use the phase plane as an analytical tool of considerable power and simplicity instead of merely regarding it as one of the very numerous ways of obtaining graphical solutions, as is the tendency in current literature. At all times, the qualitative aspects of system responses will be kept in the foreground. The aim is to obtain the greatest possible degree of insight into the problem with very little labor. For numerical answers it is best to have recourse to machine computation or to various graphical methods such as the isocline method,[1,8] the slope-line method,[9–11] or the acceleration-plane method.[12,13]

Virtually no mention will be made in this paper of such well-entrenched notions of linear servo theory as settling time, bandwidth, frequency response, etc. It is probably impossible even to define meaningfully such quantities in relation to nonlinear systems before a sufficiently broad qualitative theory is available. This paper undertakes to supply such a reasonably broad theory within the limitations of the phase-plane analysis.

The basic idea of the method is the representation of nonlinear transients as the combination of linear transients. This implies that the phase-plane trajectories of a nonlinear system are to be decomposed into segments, each of which belongs to a linear system. The pro-

---

Paper 54-519, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted January 25, 1954; made available for printing August 31, 1954.

Rudolf E. Kalman is with the E. I. du Pont de Nemours & Company, Wilmington, Del.

This work was sponsored by the U. S. Navy under Contract No. *NOrd-11799*, while the author was with the Servomechanisms Laboratory, Massachusetts Institute of Technology.

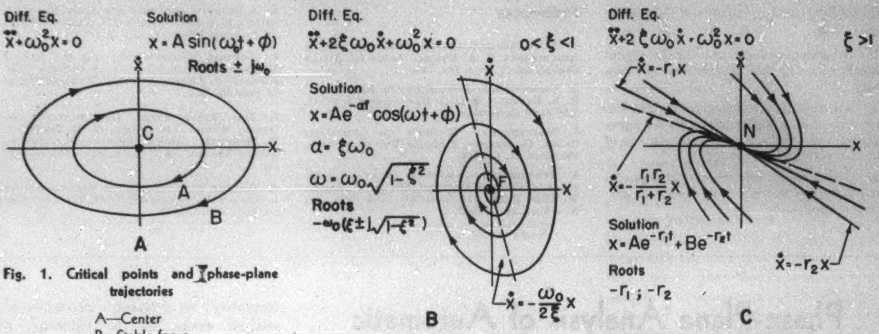

Diff. Eq.
$\ddot{x}+\omega_0^2x=0$
Solution
$x=A\sin(\omega_0t+\phi)$
Roots $\pm j\omega_0$
$\dot{x}$
x
C
A
B
A
Diff. Eq.
$\ddot{x}+2\xi\omega_0\dot{x}+\omega_0^2x=0$
$0<\xi<1$
Solution
$x=Ae^{-\alpha t}\cos(\omega t+\phi)$
$\alpha=\xi\omega_0$
$\omega=\omega_0\sqrt{1-\xi^2}$
Roots
$-\omega_0(\xi\pm j\sqrt{1-\xi^2})$
$\dot{x}$
x
F
$\dot{x}=-\frac{\omega_0}{2\xi}x$
B
Diff. Eq.
$\ddot{x}+2\zeta\omega_0\dot{x}+\omega_0^2x=0$
$\xi>1$
$\dot{x}=-r_1x$
$\dot{x}$
N
x
$\dot{x}=-\frac{r_1r_2}{r_1+r_2}x$
Solution
$x=Ae^{-r_1t}+Be^{-r_2t}$
Roots
$-r_1;-r_2$
$\dot{x}=-r_2x$
C

Fig. 1. Critical points and phase-plane trajectories

A—Center
B—Stable focus
C—Stable node

cedure will be illustrated by several examples. Although the method itself is extremely simple, its application is often quite subtle and warrants careful and detailed discussion.

The analysis requires the following major steps.

1. Approximate nonlinearities by straight-line segments.
2. Subdivide the phase plane into regions within which the trajectories are those of a linear system.
3. Assign a critical point (node, focus, etc.) to each region; this will specify all trajectories within that region.
4. Connect the trajectories belonging to various regions to obtain the transient response.

See the section entitled "Summary of the Method" for details.

An elementary knowledge of phase-plane theory is assumed. The following section gives a concise review of the mathematical facts needed.

Mathematical Review

The following is a brief summary of important properties of the phase plane which are needed for the discussion.[1,8]

UNIQUENESS

Under very general conditions, a second-order nonlinear differential equation of the type, say

$$\ddot{x}+f(x,\dot{x})\dot{x}+g(x,\dot{x})=0, \tag{1}$$

possesses a unique solution which is completely determined by the arbitrary initial conditions $x(t_0)$ and $\dot{x}(t_0)$. Essentially, it is required that the functions entering in equation 1 be analytic. From an engineering standpoint, this would exclude such double-valued functions as backlash and hysteresis; however, the phase-plane analysis can usually be modified to handle such difficulties.[4,5] All possible initial conditions of the second-order differential equation 1 may be represented as points in the phase plane $(x, \dot{x})$. It then follows that every such point will generate a solution of equation 1 which may be regarded as a parametric function of time $H[x(t), \dot{x}(t)]$. This is a curve in the phase plane, called the trajectory.

Alternately, one is led to the notion of a trajectory by considering a slightly more general version of equation 1, which is a standard form in most of the nonlinear literature

$$\begin{aligned}\frac{dx}{dt}&=P(x,\dot{x})\\ \frac{d\dot{x}}{dt}&=Q(x,\dot{x})\end{aligned} \tag{2}$$

It is easy to eliminate the time from equations 2 so that a first-order differential equation is obtained, the solution of which is the trajectory $\dot{x}(x)$

$$\frac{d\dot{x}}{dx}=\frac{\ddot{x}}{\dot{x}}=\frac{Q(x,\dot{x})}{P(x,\dot{x})} \tag{3}$$

As a consequence of the uniqueness property, there will pass only one trajectory through every ordinary point in the plane $(x, \dot{x})$. The exception is the critical point which is explained in the following.

CRITICAL (SINGULAR) POINTS

For the present purposes, it is convenient to define a critical point as a point in the phase plane where the system has stable or unstable equilibrium, i.e., where the derivatives $\dot{x}$ and $\ddot{x}$ are zero. In terms of equation 2, this means letting

$$P(x,\dot{x})=0 \text{ and } Q(x,\dot{x})=0 \tag{4}$$

The solution of equation 4 for x, $\dot{x}$ gives the location of the critical points in the phase plane. An inspection of equation 3 shows that at a critical point the slope of a trajectory becomes indeterminate; it follows that any number of trajectories may intersect at a critical point.

It is best to illustrate the various important types of critical points by means of well-known solutions of second-order linear differential equations with constant coefficients. As will be seen, the nature of the trajectories in the neighborhood of a critical point is intimately associated with the characteristic roots of these differential equations. In the case of a general nonlinear differential equation such as 2, the trajectories in the neighborhood of a critical point are given approximately by the linear perturbation equation of 2, obtained by replacing P and Q by Taylor-series expansions around the critical point and retaining only the terms linear in x and $\dot{x}$. The characteristic roots of the linear perturbation equation determine the stability properties of the critical point. In exceptional cases, when the linear terms in the Taylor-series expansions of P or Q vanish, a more elaborate theory must be used.

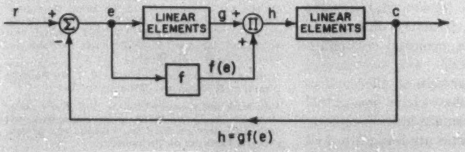


Fig. 2. General system configuration

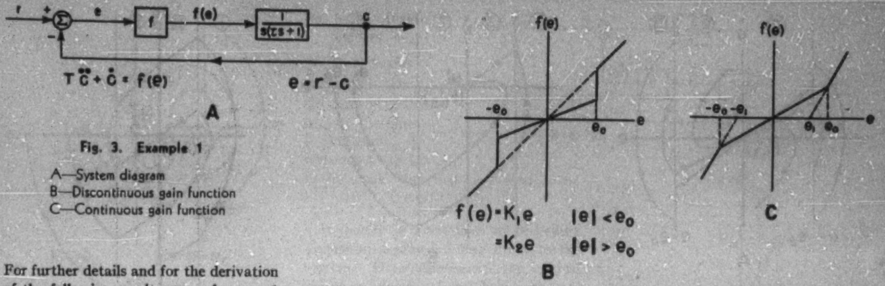


Fig. 3. Example 1

A—System diagram
B—Discontinuous gain function
C—Continuous gain function

For further details and for the derivation of the following results, see references 1 and 8.

The center (vortex) is characteristic of conservative systems; see Fig. 1(A). Its neighboring trajectories represent steady-state oscillations which are neither increasing nor decreasing. Oscillations around a center should not be confused with limit cycles which occur in dissipative nonlinear systems and represent an entirely different phenomenon. Note that a center is neither stable nor unstable; any sudden disturbance may place a point moving on trajectory *A* onto *B* and vice versa. In linear conservative systems, the trajectories around the center are concentric similar ellipses.

The focus, Fig. 1(B), is associated with a linear system whose characteristic roots are complex conjugate. Depending on whether the real part of the roots is negative or positive, the focus is stable or unstable respectively. In an oblique rectilinear co-ordinate system, the trajectories belonging to the focus are a family of logarithmic spirals.

The node, Fig. 1(C), is associated with a linear system whose characteristic equation has real roots of like sign. Depending on whether the roots are negative or positive, the node may be stable or unstable respectively. In an oblique rectilinear co-ordinate system, the trajectories belonging to the node are a family of parabolas.

There are also other types of critical points but these will not be encountered in the ensuing analysis. In Figs. 1(A), (B), and (C), the critical points happen to be at the origin; in general, they may be anywhere.

In the linear case, the trajectories belonging to a critical point are completely determined by the characteristic roots, i.e., the complex natural frequencies of the system. To simplify the terminology, the nature of a critical point will henceforth be referred to as the natural frequencies of the system. Thus, knowledge of the nature and location of a critical point is equivalent to the knowledge of all trajectories of a linear system.

It should be observed that in all the foregoing cases the trajectories rotate clockwise. The unstable trajectories may be obtained from Figs. 1(B) and (C) by reflection about the $\dot{x}$ axis; this corresponds to changing the sign of the real parts of the characteristic roots. Stable trajectories move inward, to the origin, unstable trajectories move outward, away from the origin, as the time increases in the direction of the arrows. Thus, stable (unstable) foci and nodes may be called sinks (sources) of trajectories.

Input Functions

It should be noted that equations 1 and 2 contain no forcing (time-dependent) terms. This restriction is necessary because a forcing term will, in general, increase the dimensionality of the problem. However, if the inputs are confined to steps and/or ramps ($a+bt$; a and b arbitrary, $t>t_0$), the problem may still be 2-dimensional. To see this, consider a simplified version of equation 1

$$\ddot{x}+f(x, \dot{x})\dot{x}+x=F(t)=a+bt,\ t>t_0 \tag{1A}$$

Upon substituting $y(t)=F(t)-x(t)$, equation 1(A) becomes

$$\ddot{y}+f(a+bt-y,\ b-\dot{y})(\dot{y}-b)+y=0 \tag{1B}$$

or

$$\frac{dy}{dt}=P(y, \dot{y})=\dot{y} \tag{2A}$$

$$\frac{d\dot{y}}{dt}=Q(y, \dot{y})=-[f(a+bt-y,\ b-\dot{y}(\dot{y}-b)+y]$$

Hence, either the condition $f(y, \dot{y})\equiv f(\dot{y})$, or $b=0$ must hold for the problem not to be explicitly time-dependent. It may be shown similarly that the second-order servo problem with step inputs is always 2-dimensional; with ramp inputs, this is the case only if the servo has zero steady-state position error.

Observe that the location of the critical points, given by setting $P=0$, $Q=0$ in equation 2(A), may depend on the forcing function $F(t)$. The only other information needed for constructing the transient response is the state of the system, i.e., the initial conditions, immediately after the application of the input at t_0^+. This may be determined by standard methods of linear theory.

Outline of Method

The logical first step in the study of any nonlinear system is the determination of the location and nature (node, focus, etc.) of the critical points. The local properties of the system, i.e., behavior near equilibrium points, should first be explored. This can be done without major difficulties in the case of servomechanisms. Unfortunately, no general procedure exists for investigating in-the-large properties of solutions which is the central problem of transient analysis. A fundamentally new approach is necessarily called for at this stage.

The underlying idea of this paper is the recognition that the over-all study of a wide class of nonlinear control systems becomes fairly simple whenever it is possible to subdivide the phase plane into discrete regions within which the system behaves linearly. At first, the characteristics of nonlinear elements in the loop with straight-line segments are approximated. Then, the phase plane is subdivided such that a discrete region is assigned to each line segment in the straight-line approximation. If any segment were taken by itself and extended indefinitely, a linear system would be obtained; hence, in each region the trajectories are governed by a linear differential equation. Therefore, the knowledge of the nature and location of the critical point of each region completely specifies the trajectories in that region. Knowledge of the trajectories in all the regions is equivalent to the solution of the transient problem. Actually, this is a new and extended

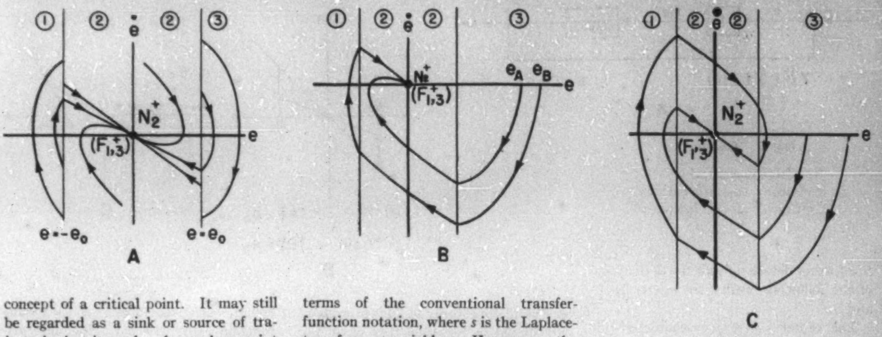


Fig. 4. Example 1, gain as in Fig. 3(B), step input

A—Phase-plane trajectories
B—Transient responses
C—Extreme trajectories for 0, 1, and 2 overshoots

concept of a critical point. It may still be regarded as a sink or source of trajectories but it need no longer be a point of equilibrium and, therefore, need not be located inside the region to which it is assigned. These matters are carefully illustrated by the examples that follow.

It follows from the foregoing arguments that the proposed method leads to an exact determination of the transient response, except for the approximation of the nonlinear element by straight-line segments. The latter qualification is unimportant since the approximation may be arbitrarily refined by increasing the number of straight-line segments; moreover, since this paper is concerned exclusively with qualitative questions, only the systems where a conveniently simple straight-line approximation has already been made will be considered.

The scope of applicability of the method embraces the class of all second-order feedback systems containing nonlinear gain elements. The term "gain element" is defined as any function generator which does not depend on frequency. A gain element may be a simple cascade element such as a saturating amplifier (whose output is $f(x)$ for input x), as in examples 1 and 2, or the more general type, illustrated in Fig. 2, where the combination of the multiplier (denoted by Π) and the box $f(e)$ may be considered as a gain element with the output given by $h=gf(e)$ at all frequencies, as in example 3. In general, gain elements may be present anywhere in the loop and may be functions of any loop variable, but only single-valued gain elements will be discussed here.

Example 1

As a first illustration, a simple positional servomechanism, Fig. 3(A), whose output element may be, for instance, an armature-controlled d-c motor is examined. (Linear components of a system will be represented for convenience in terms of the conventional transfer-function notation, where s is the Laplace-transform variable. However, the method to be outlined has no connection with the Laplace-transform theory.) The interesting feature of the system is its nonlinear (broken-line) gain element; see Fig. 3(B). It is hoped to obtain thereby a system which will be slow, or narrow-band, for small errors so as to minimize the corrupting effect of low-amplitude wide-band noise, and fast, or wide-band, for large errors. Such a system has been discussed in reference 14.

First the critical points are determined, assuming a step input ($\dot r=0$)

$$\frac{d\dot e}{de}=\frac{\ddot e}{\dot e}=-\frac{1}{\tau}\,\frac{\dot e+f(e)}{\dot e}=\frac{Q}{P},\ \text{or}\ P=0,\ Q=0 \qquad (5)$$

Hence the critical point is at $e=0$, $\dot e=0$ since $f(0)=0$, in Fig. 3(B).

By a suitable choice of the gain function $f(e)$, Fig. 3(B), it may be assumed that the closed-loop system is overdamped for $|e|<e_0$ (gain $=K_1$) and underdamped for $|e|>e_0$ (gain $=K_2$). It is then clear that the problem of determining the response to a step input is to match the solutions corresponding to different types, overdamped or underdamped, of system response as the error passes from one region into another.

In view of the foregoing assumption, the critical point would be a stable node if the gain were K_1 for all e; likewise, if the gain were K_2 for all e, the critical point would be a stable focus. In the phase plane, Fig. 4(A), this means that for $|e|<e_0$ (region 2) the trajectories converge to node N_2^+ at the origin, while for $|e|>e_0$ (regions 1 and 3) the trajectories behave as though converging to focus $(F_{1,3}^+)$ at the origin. There are really three critical points, one for each region. In the present case, all three happen to be located at the origin. Note that in each region the system behaves linearly. Subscripts of letters designating critical points refer to regions of validity. The letter is put in brackets if the critical point does not lie inside its own region ("virtual" critical point); the critical point without brackets is thus the final point of equilibrium. The superscripts + and − will denote a stable and unstable critical point respectively; trajectories converge to it as $t\rightarrow+\infty$ or $-\infty$.

The construction of the step response, at least in principle, is exceedingly simple. When a trajectory reaches the boundary $|e|=e_0$, it is merely necessary to connect it with a trajectory belonging to the critical point of the next region and going through the same point on the boundary. That this is always possible is guaranteed by the uniqueness property mentioned in the foregoing (one and only one trajectory through each ordinary point). Two such composite trajectories are shown in Fig. 4(B); they are transient responses to step inputs of magnitude e_A and e_B.

It is now apparent that the knowledge of the nature of the two critical points together with their regions of validity is all that is needed to construct a complete set of qualitative and, if necessary, quantitative step responses. For instance, one may pick out segments on the line $\dot e=0$ according to the number of times the transient will overshoot. Fig. 4(C) shows the extreme trajectories for 0, 1 and 2 overshoots.

The particular form of $f(e)$ as shown in Fig. 3(B) represents no restriction; e.g., the gain function of Fig. 3(C) may be just as easily treated, with the only

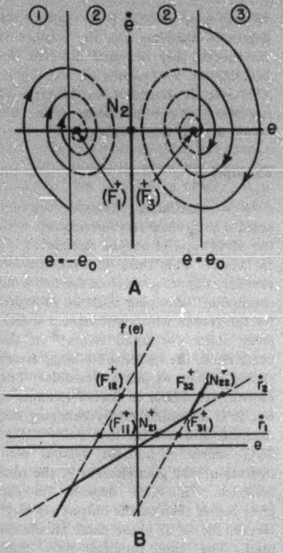


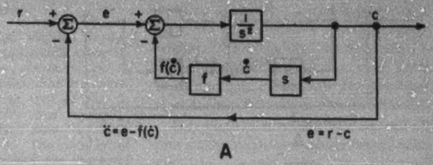
Fig. 5. Example 1, gain as in Fig. 3(C)

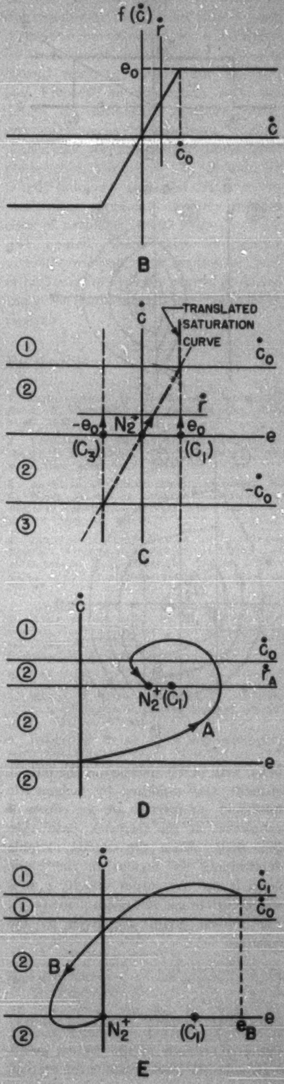
A—Phase-plane trajectories, step input
B—Determination of location of critical points for ramp input

difference that now the three critical points are separated. They are located at $\dot e=0$, $f(e)=0$. The condition $f(e)=0$ has the following significance. Two foci, (F_1^+) and (F_3^+), located at $\pm e_1$, are the intersections of the continuations of straight-line segments, corresponding to high gain and shown dotted in Fig. 3(C), with the line $f(e)=0$; node N_2^+, corresponding to the low-gain segment in Fig. 3(C), is located at the origin as before. The phase-plane situation is shown in Fig. 5(A). Extensions to other types of $f(e)$ are obvious. Even when $f(e)$ is a smooth curve, say a cubic, an approximation such as in Fig. 3(C) may be sufficiently accurate for a large operating range; if not, five or more straight-line segments may be used. The latter case may occur, for instance, when the open-loop transfer function in Fig. 3(A) is $G(s)=(s+b)/s(s+a)$ with $b>a>0$, where increasing the loop gain (i.e., increasing $df(e)/de$) leads first to a stable node, then to a stable focus, and then again to a stable node—clearly there will be five principal regions of interest. Finally, it is seen that if $f(e)$ is skew-symmetric, all the solutions in the phase plane are also skew-symmetric.

Fig. 6 (right). Example 2

A—Block diagram
B—Gain function
C—Locus of critical points ($\dot r$ increasing)
D—Trajectory for ramp input ($\dot r=\dot c_A$)
E—Trajectory for step and ramp input ($r=e_B$, $\dot r=-\dot c_B$)





It should be noted that the preceding discussion assumed a step input to the system. Determination of the critical points in equation 5 was based on $\dot r=0$. For a velocity (ramp) input, equation 5 must be modified as follows

$$\frac{d\dot e}{de}=\frac{\ddot e}{\dot e}=-\frac{1}{\tau}\,\frac{(\dot e-\dot r)+f(e)}{\dot e}=\frac{0}{0} \qquad (5\text{A})$$

$$\dot e=0,\ f(e)=\dot r$$

The critical points still lie on the line $\dot e=0$ in the phase plane; their e coordinate, however, is now determined by the intersection of the line $f(e)=\dot r$ with the extensions of the straight-line segments; see Fig. 5(B). Two different possibilities of the location of critical points depending on the value of $\dot r$ are indicated in Fig. 5(B). It is seen that for a large ramp input $\dot r_2$ the focus F_{32}^+ has become the final point of equilibrium; in other words, the transient response will be oscillatory after the error becomes sufficiently small. A velocity input thus modifies system behavior in an essential fashion; in particular, the critical points remain stationary only as long as all the inputs originate on the line $\dot r=\dot e=$ constant. The response of the system to inputs on this line will further differ according to initial error (number of overshoots, etc.). Needless to say, once the critical points and regions are known, the trajectories are matched on the boundaries in the same manner as in Fig. 4(B) and 4(C).

For any step or ramp input or a succession of these, knowledge of the behavior of critical points supplies all the necessary qualitative information to guide the design procedure or the actual computation of the response. The circumspection gained through the introduction of the generalized critical-point concept renders the foregoing analysis far more powerful and simple than the algebraic method employed in the original treatment of this system by Beck and Blumenthal.[14]

Example 2

This is the investigation of another simple positional servomechanism, Fig.

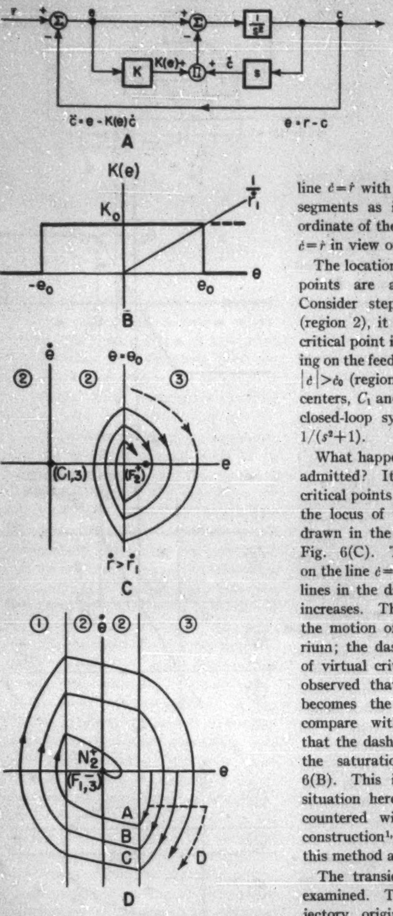


Fig. 7. Example 3

A—Block diagram
B—Gain function
C—Transient response to large ramp ($\dot r > \dot r_1$)
D—Appearance of unstable limit cycle ($\dot r = 0$)

6(A), with $(1/s^2)$ (inertia) as the output element, and stabilized by tachometric feedback. Of interest are the effects of saturation in the feedback path; see Fig. 6(B). Since the crucial variable determining the regions in the phase plane is now the output velocity $\dot c$, it is convenient to use the phase plane $(e, \dot c)$. The critical points are given by the equation

$$\frac{d\dot c}{de}=\frac{\ddot c}{\dot e}=\frac{e-f(\dot c)}{\dot e}=\frac{0}{0} \qquad (6)$$

$$\dot e=0,\ e=f(\dot r-\dot e)=f(\dot r)$$

The e co-ordinate of the critical points is given by the intersection of the vertical line $\dot c=\dot r$ with the extended straight-line segments as in Fig. 6(B). The $\dot c$ co-ordinate of the critical points is given by $\dot c=\dot r$ in view of $\dot e=0$.

The location and nature of the critical points are again easy to establish. Consider step inputs first. If $|\dot c|<\dot c_0$ (region 2), it may be assumed that the critical point is a node (or focus, depending on the feedback gain) N_2^+; whereas if $|\dot c|>\dot c_0$ (regions 1 and 3), there are two centers, C_1 and C_3, since in that case the closed-loop system transfer function is $1/(s^2+1)$.

What happens if ramp inputs are also admitted? It is easy to see that the critical points will move around; in fact, the locus of the critical points can be drawn in the phase plane, as shown in Fig. 6(C). The critical points start on the line $\dot c=0$ and will move on straight lines in the direction of the arrows as $\dot r$ increases. The dash-dotted line indicates the motion of the final point of equilibrium; the dashed lines show the motion of virtual critical points. It should be observed that for $\dot r>\dot c_0$, the center C_1 becomes the final equilibrium point; compare with Fig. 5(B). It is seen that the dash-dotted line is the same as the saturation characteristic in Fig. 6(B). This is not an accident. The situation here is the same as that encountered with the so-called Liénard construction[1,7] which is contained in this method as a special case.

The transient response may now be examined. The point at which a trajectory originates corresponds to the initial velocity of the output element and the initial error caused by the input; the trajectory will tend to a critical point determined by the input velocity. Moreover, since $c=-e+\text{constant}$, all the trajectories will be winding in the counterclockwise direction. This is merely a consequence of the particular choice of phase-plane variables. Two typical cases are the following: In Fig. 6(D), trajectory A is the response from rest to a ramp input $\dot r_A$; the motion starts at the origin and tends to the node N_2^+ as shown. In Fig. 6(E), trajectory B results from a step input $r=e_B$ and a ramp input $\dot r=-\dot c_B$ at the moment when the steady-state output velocity is $\dot c_B$ $(e=0)$. It is seen that the degree of damping, i.e., the number of overshoots, may be quite different depending on the nature of the input. This over-all picture of specific types of response is not available from the describing-function method.[15,16]

Example 3

As a concluding example, consider again a positional servomechanism with the same type of output element $(1/s^2)$ as before. This time, the stabilization problem is approached from the following standpoint: damping must be provided for the system which will make it sufficiently stable for small errors; on the other hand, the response for large errors should remain as fast as possible. The objective is similar to that of example 1, but it is accomplished by changing the damping and not the loop gain.[17] Fig. 7(A) shows a possible system configuration; the gain element of the rate feedback, Fig. 7(B), depends on the error rather than on the output velocity. Hence, the $(e, \dot e)$ phase plane should be used. The gain function $K(e)$ may be obtained, for instance, by using a relay. The fact that $K(e)$ is discontinuous (staircase function) is not accidental; any other function $K(e)$ would have to be approximated in this way because of the multiplication in the loop.

The critical points are obtained from

$$\frac{d\dot e}{de}=\frac{\ddot e}{\dot e}=-\frac{e+K(e)(\dot e-\dot r)}{\dot e}=\frac{0}{0} \qquad (7)$$

$$\dot e=0,\ e=K(e)\dot r$$

When $\dot r=0$, it is seen by inspection that there are again three regions; the critical point corresponding to the inner region is a focus F_2^+ (or node, depending on K_0), the critical point corresponding to the outside regions is a center $(C_{1,3})$. All critical points are located at the origin.

For $\dot r\neq 0$, the e co-ordinate of the critical points is given by the intersection of the line $(1/\dot r)e$ with discontinuous horizontal segments of $K(e)$; see Fig. 7(B). The intersection with $K(e)=K_0$ determines the location of the node while the centers are located at the intersection with $K(e)=0$, i.e., they remain at the origin irrespective of the value of $\dot r$. Now if the input velocity increases beyond a critical value ($\dot r_1$ in the figure), the node will always lie in the region $|e|>e_0$ and neither critical point will lie inside its own region. Is there a final point of equilibrium?

Consider an arbitrary trajectory, as shown in Fig. 7(C), composed of arcs belonging to the center $C_{1,3}$ and the focus (F_2^+). Such a trajectory must evidently converge to the point e_0 as shown. The same situation occurs when the second critical point is a node (N_2^+). The point e_0 is clearly not a critical point; yet it may be said to be stable since for any perturbation the trajectories will converge to it, assuming that the condition $\dot{r}>\dot{r}_1$ remains in effect. It is not possible to show by presently known methods other than those using phase-plane concepts that e_0 will represent the steady-state position error of the system for all $\dot{r}>\dot{r}_1$ and that it has the stability property pointed out in the foregoing. The damping around e_0 will be very poor (in region 3, zero) regardless of whether the critical point in region 2 is a focus or node and, therefore, the position error will be oscillatory near the equilibrium point. An analogue computer study of this system[17] leads to similar conclusions.

It is instructive to inquire what happens when the damping becomes negative instead of zero for $|e|>e_0$. It may be assumed that the critical points will be a stable node N_2^+ and an unstable focus ($F_{1,3}^-$) respectively; the inputs are to be restricted to steps. The situation is shown in Fig. 7(D). It follows from symmetry considerations that it is possible to have one and only one closed trajectory, B, which is called a limit cycle; examination of the neighboring trajectories shows that for smaller steps, A, they converge to the origin and that for larger steps, C, they diverge and become unstable. Therefore, the limit cycle B is unstable. It represents the boundary between stable and unstable regions in the phase plane.

Care must be taken to interpret properly the domain of stability, which is the region inside the limit cycle. For example, if the inputs are assumed to be steps, not merely their amplitude but also their time of occurrence must be taken into account. To see this, observe that, if a step input corresponding to A is followed by another step input of the same polarity before the first transient is substantially over, the system may become unstable, as indicated by dotted lines, D. However, the system may remain stable if the step transients are sufficiently separated in time. It would be very difficult to interpret correctly such phenomena without recourse to phase-plane concepts.

Summary of Method

The basic steps in the technique of analysis presented in the foregoing may be summarized as follows.

1. Each nonlinear gain element is approximated. If it is a cascade element, examples 1 and 2, the approximation is done by continuous straight-line segments. If it is the more general type of Fig. 2, horizontal discontinuous line segments (staircase approximation) must be used as shown in example 3. The number of line segments and, therefore, the degree of approximation is arbitrary; for a qualitative analysis, the simplest approximation which preserves the essential shape of the nonlinear characteristic is preferable.

2. The phase plane is subdivided into regions so that only one region is assigned to each line segment obtained in step 1. In doing this, a set of phase-plane variables, say x_1 and x_2, is chosen so as to make the regions as simple as possible. It should be noted that x_1 and x_2 must represent either the proper initial conditions of the system or a nonsingular linear transformation of these, since otherwise the plane (x_1, x_2) is not a phase plane. The natural choice is $x_1=e$, $x_2=\dot{e}$ but sometimes another set is more convenient; see example 2.

3. The critical points of the differential equations formulated in terms of x_1 and x_2 in step 2 are found by setting $\dot{x}_1$ and $\dot{x}_2$ to zero. Care must be taken to include the input functions (step, ramp) in determining the location of the critical points. Only one critical point is assigned to each region. The nature of a critical point in each region is obtained by solving for the characteristic roots of the linear differential equation pertaining to that region. Usually, a qualitative answer (node, focus, center) is all that is required.

4. The actual transient solution is obtained by matching on the boundary trajectories belonging to various regions. A qualitative solution may be easily sketched. The totality of the behavior of the system may be visualized by observing how the critical points, and possibly the regions, vary with the input.

5. For a quantitative solution, one would need standard plots corresponding to various possible roots of the characteristic equation. On these plots, the time may be indicated as a parameter; there are also a number of alternate methods[2,3] for obtaining values of time. It is felt, however, that, for strictly quantitative studies, the method presented here is probably inferior to numerous other graphical methods which are in existence.

The limitations of the method as presented here are:

1. Only second-order systems may be treated in the phase plane.

2. The method is exact only within the idealization made in step 1.

3. The nonlinear gain elements were restricted to be single-valued.

Conclusions

The foregoing examples were used to illustrate a generalized method of phase-plane analysis. Instead of attacking each problem separately, a systematic procedure is now available for nonlinear studies. This enables one to explore a given problem analytically with little effort and gain insight into the principal mechanisms of nonlinear system behavior.

This method is believed to contain all previously published phase-plane analyses as special cases. Its principal shortcoming at the present time is that it cannot be extended in a straightforward fashion to high-order systems. Recent work by Bogner and Kazda[18] and Preston[19] in connection with the design of third-order optimum relay servomechanisms demonstrate various possibilities for attacking 3-dimensional phase-space problems. It should be noted, however, that the relay servomechanism represents a very special class of problems; aside from the work just quoted, there exist at present virtually no published studies of third or higher order dynamic control problems using the phase-space (exact) method of analysis.

References

1. Introduction to Nonlinear Mechanics (book), **N. Minorsky.** J. W. Edwards, Ann Arbor, Mich., 1947.

2. Fundamental Theory of Servomechanisms (book), **L. A. MacColl.** D. Van Nostrand Company, New York, N. Y., 1945.

3. Analysis of Relay Servomechanisms, **H. K. Weiss.** *Journal of Aeronautical Sciences*, Institute of Aeronautical Sciences, New York, N. Y., vol. 13, 1946, pp. 364–76.

4. Instrument Inaccuracies in Feedback Control Systems with Particular Reference to Backlash, **H. T. Marcy, M. Yachter, J. Zauderer.** *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 778–88.

5. Automatic and Manual Control (book), **A. Tustin,** editor. Buttersworths Scientific Publications, London, England, "An Introduction to the Analysis of Nonlinear Closed-Cycle Control Systems," **W. E. Scott,** 1952, pp. 249–62.

6. The Effect of Friction on the Behavior of Servomechanisms at Creep Speeds, **J. G. L. Michel, A. Porter.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 98, pt. II, 1951, pp. 297–311.

7. Operating Modes of a Servomechanism with Nonlinear Friction, **H. Lauer.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 255, 1953, pp. 497–511.

8. Theory of Oscillations (book), (English Translation), **A. A. Andronow, C. E. Chaikin.** Princeton University Press, Princeton, N. J., 1949.

9. Du Coup de Bélier en Hydraulique au Coup de Foudre en Électricité (book), **L. Bergeron.** Dunod, Paris, France, 1950.

10. Methods and Results from M.I.T. Studies in Unsteady Flow, **H. M. Paynter.** *Journal*, Boston Society of Civil Engineers, Boston, Mass., vol. 39, 1952, pp. 120–165.

11. A Graphical Analysis for Nonlinear Systems, **P.-S. Hsia.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 99, pt. II, 1952, pp. 125–34.

12. Nonlinear Analysis of Electro-Mechanical Problems, **Y. H. Ku.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 255, 1953, pp. 9–31.

13. A Method for Solving Third and Higher

Order Nonlinear Differential Equations, **Y. H. Ku.** *Ibid.*, vol. 256, 1953, pp. 229–44.

14. Proceedings of the First National Congress of Applied Mechanics (book). American Society of Mechanical Engineers, New York, N. Y., "Transient Analysis of Nonlinearized Single-Lag Servomechanisms," **I. S. Blumenthal, F. J. Beck,** 1951, pp. 155–60.

15. A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms, **R. J. Kochenburger.** *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 270–84.

16. Sinusoidal Analysis of Feedback-Control Systems Containing Nonlinear Elements, **E. C. Johnson.** *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 169–81.

17. A Study of Automatic Control Systems with Nonlinear Rate Feedback, **Robert W. Wedan.** *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1952.

18. An Investigation of Switching Criteria for Higher-Order Contactor Servomechanisms, **Irving Bogner, Louis F. Kazda.** *AIEE Transactions*, vol. 73, pt II, July 1954, pp. 118–27.

19. Nonlinear Control of a Saturating Third-Order Servomechanism, **J. L. Preston.** *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1954. Also *Technical Memorandum No. 6897–14*, Servomechanisms Laboratory.

The Nature of Voltage Ripple on D-C Generators

F. N COLLAMORE
ASSOCIATE MEMBER AIEE

THE output voltage of a d-c generator contains an a-c component which is superimposed upon the d-c output of the generator. This voltage component, or voltage ripple, may be less than 1 per cent of the output voltage or may have peaks as high as 10 per cent of the d-c output, depending upon the design and condition of the generator. For many applications of d-c power the voltage ripple is of no consequence and may be neglected; however, there are other applications, such as some types of electronic equipment where d-c power of high purity is essential. In these cases the voltage ripple must be reduced to a low value either by design of the generator, filtering of the output, or both. Where the voltage ripple is to be reduced by filtering, it is desirable to know the magnitude and frequency of the various components that make up the voltage ripple. If the ripple is to be reduced by design of the generator, it is necessary to know the cause of each of the ripple-voltage components.

The data presented in this paper were obtained from tests on various types of generators used on aircraft which, by their nature as high-speed, high-output machines, may exhibit a higher percentage of voltage ripple than more conservatively designed industrial type of generators. However, the causes and methods of reduction of voltage ripple should apply to any type of d-c generator.

To determine the characteristics of the voltage ripple in the output of a d-c generator, the magnitude and frequency of each of the ripple-voltage components should be obtained and the magnitude of the combined ripple-voltage wave determined. It is also necessary to know the voltage-ripple wave shape to determine the causes of the ripple. The test setup used to obtain the data presented is shown in Fig. 1. A 4-microfarad condenser is used in series with the measuring instruments to isolate the voltage ripple from the direct voltage. The wave analyzer is used to determine the magnitude of each ripple frequency while the calibrated oscilloscope is used to measure the peak-to-peak values of the total voltage ripple and to show its over-all wave shape. A cathode-ray oscilloscope with a high-frequency response is necessary to obtain an accurate picture of the voltage ripple because some of the ripple frequencies are very high. By using a suitable cathode-ray tube, photographs may be taken of the voltage wave. It was found that, because of the irregularity of some of the ripple voltages, it was necessary to use the single-sweep feature of the oscilloscope to obtain good photographs without multiple exposures.

The voltage ripple found in the output of most d-c generators is usually the result of several different causes which, when combined, result in a very irregular voltage wave shape. The major causes are:

Cause of Voltage Ripple

1. Variations in field excitation.
2. Variations in speed.
3. Improper physical alignment.
4. Variable reluctance in armature laminations caused by grain orientation.
5. Variable air-gap and armature-tooth reluctance caused by armature slots.
6. Commutator bar ripple.
7. Brush bounce caused by commutator roughness.

Frequency

1. Same frequency as field current variation.
2. Same frequency as speed variations.
3. Revolutions per second; and multiple thereof.
4. Revolutions per second × poles and multiples thereof.
5. Revolutions per second × number of armature slots and multiples thereof.
6. Revolutions per second × number of commutator bars × denominator of number of commutator bars between brushes reduced to its lowest terms.
7. No set frequency. Voltage fluctuations occur whenever brush-arcing occurs.

It can be seen that there are three general categories into which the ripple may be divided: variations in generator input; generator design and manufacturing tolerances; and generator deterioration with age. For the most part, this paper is concerned with the voltage resulting from design and manufacturing tolerances which are items 3, 4, 5, and 6.

It is known that the sheet steel used for

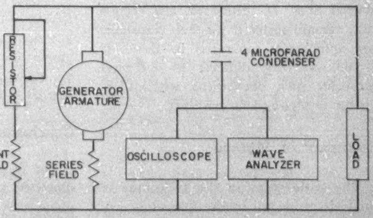


Fig. 1. Test setup

Paper 54-420, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 29, 1954; made available for printing August 30, 1954.

F. N. Collamore is with Jack & Heintz, Inc., Cleveland, Ohio.

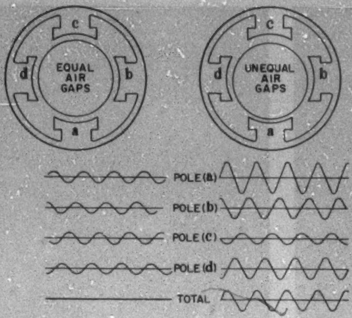


Fig. 2 (left). Slot voltage ripple

Fig. 3 (right). Voltage ripple on a 250-volt d-c generator

Top—No load
Center—Full load
Bottom—Calibration 1,000 cycles per second, 10.0 volts peak to peak

armature laminations has better magnetic properties in one direction than the other. This is because of grain orientation resulting from rolling the steel which gives the steel better magnetic properties in the direction of rolling than at right angles to it. If the armature-core laminations are assembled with the grain in one direction, there will be a variation in tooth and core reluctance at different positions around the armature which will result in flux pulsations as the armature rotates. These flux pulsations will result in a voltage ripple of 1 cycle per pole per revolution of the armature. However, as these flux pulsations are not necessarily sinusoidal in nature, there may be ripple frequencies at multiples of the basic frequency. The voltage ripple resulting from grain orientation may be minimized by stacking the armature laminations in such a manner that the grain orientation is distributed in several directions. This will reduce the ripple at the basic frequency but may increase it at some of the higher frequencies where it may be more easily filtered.

The effect of variations in the air-gap and armature-tooth reluctance as a result of the armature slots causes a flux pulsation resulting in a voltage ripple. The effect of the slot ripple becomes increasingly more pronounced as the ratio of slot width to air gap at the pole tip becomes larger. This is of particular importance in aircraft generators where small air gaps are used. Each time an armature slot passes under a main pole, a variation in air-gap and tooth reluctance occurs resulting in a cycle of slot ripple voltage. The basic slot voltage-ripple frequency is 1 cycle per slot per revolution of the armature. Inasmuch as the slot voltage ripple usually is not sinusoidal, there will be harmonics of the slot frequency which show up as multiples of the basic slot frequency. The magnitude of the slot voltage ripple is considerably influenced by the design and manufacturing tolerances of the generator.

If a generator has a fractional number of armature slots per pole pitch, theoretically the basic slot ripple will cancel out. As an example consider a 4-pole wave-wound generator with $N \pm {}^1/_4$ slots per pole pitch and equal air gaps, as illustrated in Fig. 2. Inasmuch as 1 cycle of ripple voltage is caused each time a slot passes a main pole tip, and the relative position of slots to pole tips is displaced $^1/_4$ slot between adjacent poles, it can be seen that the ripple voltage caused by one pole will be 90 degrees out of phase with the ripple caused by adjacent poles. The ripple voltage generated by opposite poles is 180 degrees out of phase and therefore, cancels out. However, harmonics of the basic frequency will not necessarily cancel out, and in some cases, such as the fourth harmonic in this example will add together in the output. Any mechanical misalignment such as improper pole spacing will tend to prevent the basic ripple-voltage frequency from being completely cancelled.

In many cases, because of manufacturing tolerances in building generators, the air gaps will not all be equal. An example of this is shown in Fig. 2. If this is also a wave-wound generator with a fractional number of slots per pole pitch, the basic slot ripple will not be completely cancelled. The magnitude of the slot ripple is dependent upon the ratio of slot width to air-gap length which causes the poles with the small air gaps to produce the higher slot ripple. This results in a ripple voltage in the output of the generator at slot frequency, with the harmonic frequencies being evident at multiples of the slot frequency.

Another example of slot ripple caused by manufacturing variations is evident in a machine with an eccentric armature whose center of rotation is not the center of the main poles. This condition causes the main pole air gap at any pole to change from a minimum gap to a maximum gap with each revolution of the armature. An example of this may be an armature whose center of rotation is displaced from the center of the main poles by the same amount that the armature core is off center from the center of the shaft. This will result in equal air gaps for one position of the armature, but when the armature is rotated $^1/_2$ revolution, the air gaps become unequal. It can be seen from Fig. 2 that the voltage ripple will then change from a condition of equal air gaps, where the basic slot ripple cancels, to a condition of unequal air gaps where it does not cancel. This results in a voltage ripple which is modulated in both magnitude and frequency content at rotative frequency, as shown in Fig. 3. This modulation may be reduced by minimizing the amount of armature eccentricity. The slot ripple of a generator having a fractional number of slots per pole pitch may be reduced considerably by making the main pole air gaps equal; however, for machines with an integral number of slots per pole pitch, other means such as increasing the air gaps at the pole tips will be required.

Commutator bar ripple is the voltage ripple resulting from the action of the brushes at the commutator. The primary cause of this voltage ripple is arcing at the brushes resulting in a voltage drop from the brush to the commutator which causes a drop in the output voltage of the generator. One reason for commutator ripple is poor commutation caused by improper interpole compensation. Usually, the commutator ripple begins to appear as soon as there is visible sparking at the brushes and becomes progressively worse

Fig. 4. Effect of commutation on voltage ripple

Top—Sparkless commutation
Center—Sparking at the brushes
Bottom—Calibration 1,000 cycles per second, 3.0 volts peak to peak

Fig. 5. Voltage ripple on a 30-volt d-c generator

Top—No load
Center—Full load
Bottom—Calibration 1,000 cycles per second, 3.0 volts peak to peak

Fig. 6. Voltage ripple on a 250-volt d-c generator

Top—No load (expanded scale)
Center—Full load (expanded scale)
Bottom—Calibration 1,000 cycles per second, 10.0 volts peak to peak

with commutation as the load increases. During normal commutation of a d-c machine, the current in each armature coil is reversed as the commutator bars to which the coil is connected pass under the brushes. If, as in the case of a poorly compensated machine, the current in the coil is not completely reversed when the bar leaves the trailing edge of the brush, arcing will result. The voltage drop across this arc is reflected back in the output of the generator as ripple voltage. This will appear as a negative voltage with 1 cycle appearing each time a commutator bar passes from under a brush. Whenever there are an integral number of commutator bars between brushes, the commutator ripple frequency will be equal to the number of bars multiplied by the rotative speed in revolution per second as the bars under each brush are in the same relative position. However, in the case of a machine with a fractional number of bars between brushes, the ripple frequency must be multiplied by the denominator of the number of bars between brushes reduced to its lowest terms. This is because the position of the bars under the brushes are out of phase and do not pass from under the bars at the same time; this results in a higher frequency ripple of lower magnitude.

To show more clearly the effect of commutation upon voltage ripple, oscillograms were taken of the output voltage of a d-c generator, as the commutation was changed, by changing the interpole strength. For this test, a 6-pole lap-wound generator was used, with 57 bars and 57 slots on the armature. It was operated at 4,000 rpm giving a basic slot frequency of 3,800 cycles per second. However, as there are $9^{1}/_{2}$ slots per pole, the basic slot frequency is substantially cancelled out leaving 7,600 cycles per second, as the major slot frequency. The commutator bar ripple frequency will also be 7,600 cycles per second as the denominator of number of bars per brush is 2. The effect of commutation on voltage ripple may be clearly seen in Fig. 4. The top trace was made with no sparking at the brushes and shows a voltage ripple caused by the armature slots. The center trace shows commutator ripple resulting from sparking at the brushes. To obtain this trace, the generator interpole current was shunted until visible sparking at the brushes was obtained. A calibration trace of 1,000 cycles per second and 3.0 volts peak to peak is shown on the lower trace. These data show very clearly that poor commutation can result in a considerable increase in voltage ripple on a d-c generator.

Table I. Wave Analysis of Voltage Ripple on a 30-Volt D-C Generator

Generator Data: 6-Pole Lap-Wound, 57 Armature Slots, 57 Commutator Bars, Test Speed 4,000 rpm, Oscillograms per Fig. 5

Frequency	RMS Volts, No Load	RMS Volts, Full Load
400	0.039	0.062
1,200		0.022
2,000		0.024
2,400		0.110
2,800		0.026
3,600	0.031	0.020
3,800	0.039	0.023
4,000	0.017	0.020
4,400		0.011
5,200		0.011
7,600	0.220	0.165
15,200	0.018	0.014

Values of ripple voltage below 0.01 volt rms not recorded.

Figs. 5 and 6 show oscillograms for voltage ripple on a 30-volt and 250-volt d-c generator respectively. Tables I and II show the wave analyses of voltage ripple on a 30-volt and 250-volt d-c generator respectively.

Another cause of sparking at the brushes may be poor brush-riding quali-

Table II. Wave Analysis of Voltage Ripple on a 250-Volt D-C Generator

Generator Data: 4-Pole Wave-Wound, 21 Armature Slots, 63 Commutator Bars, Test Speed 12,000 rpm, Oscillograms per Figs. 3 and 6

Frequency	RMS Volts, No Load	RMS Volts, Full Load
200	0.11	0.12
1,600	0.22	0.16
2,400	0.20	
4,000	1.3	0.48
4,200	1.9	0.60
4,800		0.12
8,400	0.54	0.38
8,800	0.25	
12,600	1.0	0.30
12,800	0.9	0.30
16,800	1.6	0.64

Values of ripple voltage below 0.1 volt rms not recorded.

ties because of a rough commutator surface. The ripple from this cause is not usually consistent in magnitude or at any set frequency as it depends upon the brush contact with the commutator. Whenever a brush is disturbed so as to cause an arc, a voltage ripple results, the magnitude of which depends on the severity of the brush disturbance. This type of voltage ripple is more likely to develop after a generator has been in service for some time and commutator roughness has developed.

It can be seen that voltage ripple is the result of many factors in the design and building of the machine and also some factors external to the machine itself. Such factors as field voltage and speed variations are directly dependent upon the power supply to the generator. In some cases, environmental conditions, such as vibration that may cause poor brush-riding qualities, may also result in voltage variations. The design of the generator and the tolerances required in its manufacture must be carefully considered in order to hold the voltage ripple to a minimum. The armature winding, slots, poles, and air gap must be designed to minimize the generation of voltage ripple. The manufacturing tolerances which effect the ripple-voltage characteristics must also be held to a practical minimum.

The problem of building a power system with voltage ripple within given limits is one where the generator capabilities and the possibilities of filtering must be carefully evaluated. While voltage ripple may be held within certain limits by design with little penalty in cost and size, it may be true that a further reduction may be more easily accomplished by the use of a suitable ripple-voltage filter.

In either case, the ripple-voltage frequencies and their causes must be known so as to determine judiciously the requirements of a d-c power system. Careful consideration must be given as to which ripple voltages and frequencies are more easily reduced at the generator or by external filtering.

This problem is of particular importance in the case of air-borne equipment where weight and space are severely limited. As aircraft power systems become larger with more loads which have low ripple-voltage requirements, the nature of the ripple voltage and its causes will become increasingly more important.

Reference

1. Voltage Irregularities in D-C Generators, J. T. Fetch. *AIEE Transactions*, vol. 49, 1930, pp. 1068–85.

Jitter in Instrument Servos

R. L. HOVIOUS
NONMEMBER AIEE

Synopsis: Jitter in electromechanical instrument servos is studied analytically and by analogue-computer simulation. The study identifies some unreported causes of jitter, identifies the jitter-causing nonlinearities that can be eliminated, establishes relations between the unavoidable nonlinearities that will eliminate jitter, and expresses the effects of granular or quantizing elements mathematically and proves the validity of these equations. The principles are used to modify a commercial servo.

VARIOUS nonlinear elements and system-generated noises cause or contribute to jitter in electromechanical instrument servos. Jitter is unwanted noise of a predominantly oscillatory nature; it can be wide-band or can contain only one or two frequencies. Analysis by analytical methods alone is not practical because jitter has many origins and takes many forms. However, analysis by analogue-computer simulation along with application of the describing function permits successful investigation. Details of the circuits used in the simulation are discussed elsewhere.[1]

The method used in this study included:

1. Simulation of a suppressed-carrier type of servo including the important nonlinearities and noises.
2. Application of the describing function technique to supplement and guide the GEDA* analogue-computer simulation.
3. Measurements made on a specific servo.

Fig. 1 shows the block diagram of the servo and Fig. 2 shows the block diagram on which the simulation was based. The system nonlinearities included coulomb friction between the potentiometer winding and its wiper arm, and between the motor shaft and its bearings. In addition, the gear-compliance characteristic including backlash, the quantizing effect of the wire-wound potentiometer, and the saturating characteristic of the servo amplifier were considered.

Throughout the study a single nonlinear simulation based on analytical study and test data was employed for the 2-phase a-c servomotor. Fig. 3 contrasts the simulated motor torque characteristic with test data. The effects of motor nonlinearities were slight. Amplifier saturation caused acceleration limiting of the servomotor. For step-command inputs greater than approximately 5 per cent full scale, the servo was velocity-limited.

System-generated noises included the 60-cycle voltage coupled by mutual inductance from the reference or field winding of the 2-phase motor into the control winding. The most important system-generated noise was caused by the static d-c unbalance between the tubes in the push-pull output stage of the servo amplifier; this noise signal flowed through the Transicoil motor control winding. Interaction with the 60-cycle reference winding voltage caused a strong 60-cycle torque which produced a 60-cycle servo output voltage at zero-input signal. The motor and gear inertia and viscous friction were not sufficient to damp out all this torque. Also studied were the sharp spikes of noise caused by variation of the potentiometer contact resistance. The effects of the combined nonlinearities were studied, and the effects of individual non-

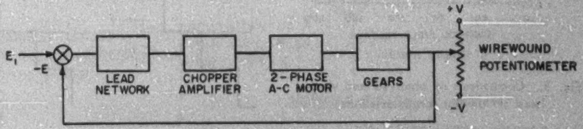


Fig. 1. Block diagram of a suppressed-carrier instrument servomechanism

---

*Trademark, Goodyear Aircraft Corporation.

Paper 54-521, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted April 20, 1954; made available for printing September 28, 1954.

R. L. Hovious is with the Goodyear Aircraft Corporation, Akron, Ohio.

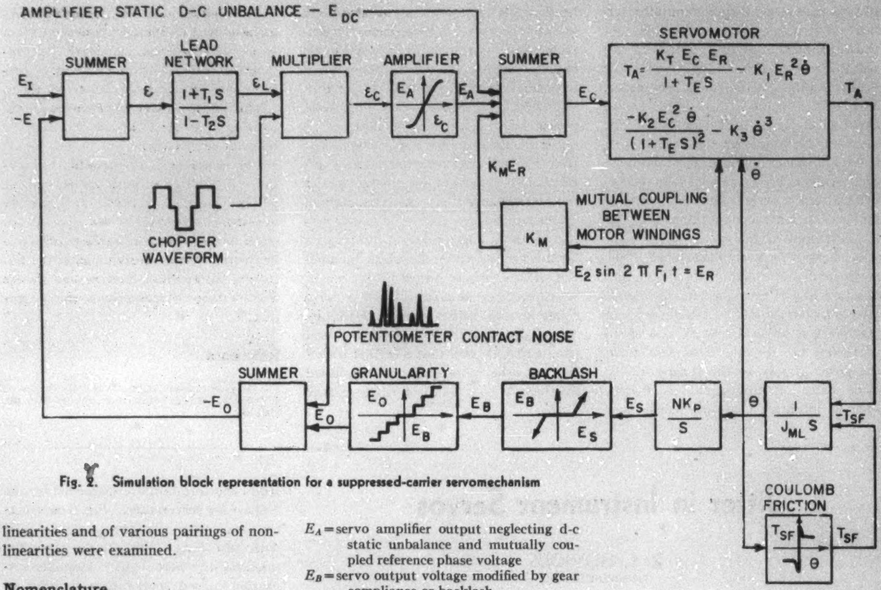

AMPLIFIER STATIC D-C UNBALANCE – E_{DC}
SERVOMOTOR
LEAD NETWORK
SUMMER
MULTIPLIER
AMPLIFIER
SUMMER
E_I
$-E$
ϵ
$\frac{1+T_1S}{1-T_2S}$
ϵ_L
ϵ_C
E_A
ϵ_C
E_A
E_C
$T_A = \frac{K_T E_C E_R}{1+T_E S} - K_1 E_R^2 \dot\theta$
$\frac{-K_2 E_C^2 \dot\theta}{(1+T_E S)^2} - K_3 \dot\theta^3$
T_A
$K_M E_R$
$\dot\theta$
CHOPPER WAVEFORM
MUTUAL COUPLING BETWEEN MOTOR WINDINGS
K_M
$E_2 \sin 2\pi F_1 t = E_R$
POTENTIOMETER CONTACT NOISE
SUMMER
GRANULARITY
BACKLASH
$-E_O$
E_O
E_O
E_B
E_B
E_B
E_S
E_S
$\frac{NK_P}{S}$
θ
$\frac{1}{J_{ML}S}$
$-T_{SF}$
COULOMB FRICTION
T_{SF}
θ
T_{SF}

Fig. 2. Simulation block representation for a suppressed-carrier servomechanism

linearities and of various pairings of nonlinearities were examined.

Nomenclature

A = nondimensional describing function
B = backlash at motor shaft
E = servo output voltage
ϵ_L = servo error modified by lead network
ϵ = servo error
ϵ_C = chopped servo error or vibrator output
E_A = servo amplifier output neglecting d-c static unbalance and mutually coupled reference phase voltage
E_B = servo output voltage modified by gear compliance or backlash
E_C = voltage input to the motor control windings
E_{DC} = static amplifier d-c unbalance
E_I = servo input voltage
E_O = servo output voltage modified by gear compliance and granularity
E_S = servo output voltage neglecting gear compliance, potentiometer granularity, and contact resistance noises
F_1 = chopper and motor reference phase frequency
J_{ML} = moment of inertia of the motor rotor, potentiometer pointer, gear train, and couplings
K_A = amplifier gain
K_M = coefficient of mutual inductance between motor reference and control windings
K_P = potentiometer sensitivity, volts per degree
K_T, K_1, K_2, K_3 = characteristic constants of the motor
N = gear train ratio
n = maximum sinusoidal input excursion; see Fig. 6
Q = volts per step in the quantizing element
S = Laplace operator
T_A = torque at the motor output shaft
T_{av} = average value of torque at the motor output shaft
T_E = electric time constant of the motor input winding
T_{SF} = coulomb friction expressed as a torque at the motor output shaft
T_1, T_2 = time constants of the lead network
$\dot\theta$ = motor velocity
ω = servomotor angular velocity, revolutions per minute

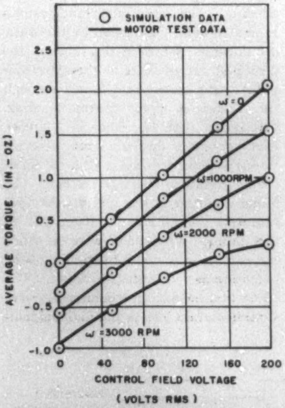


Fig. 3. Comparison of observed and simulated servomotor torque relations

$T_{av} = 0.0104E_C - 0.255\times10^{-3}\omega - 0.7\times 10^{-6}\omega E_C^2 - 10^{-11}\omega^3$

Granularity Effects

The describing function previously has been used to extend the frequency-response method of analysis to servomechanisms containing nonlinearities such as limiting, dead space, and coulomb friction.[2-4] In this study, this technique was applied to systems containing granu-

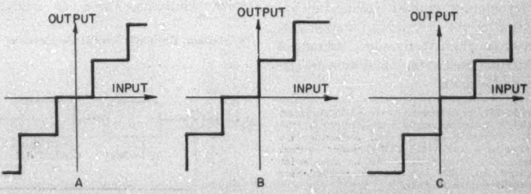

OUTPUT
INPUT
A
OUTPUT
INPUT
B
OUTPUT
INPUT
C

Fig. 4. Input-output relations for a granular device

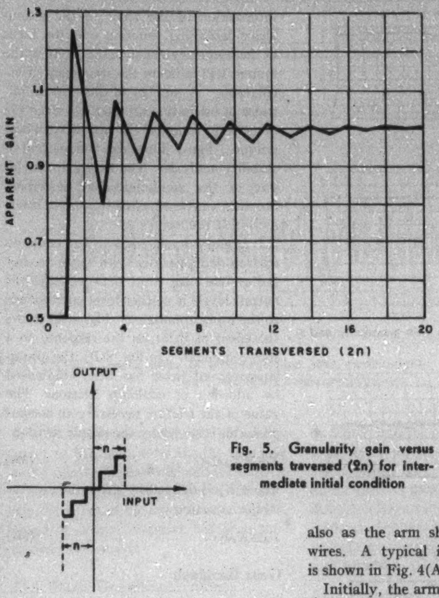


Fig. 5. Granularity gain versus segments traversed (2n) for intermediate initial condition

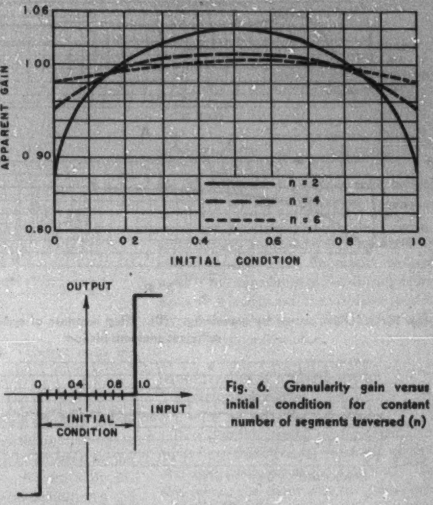


Fig. 6. Granularity gain versus initial condition for constant number of segments traversed (n)

lar elements. The conclusions were confirmed by analogue-computer simulation.

A simplified analysis was made of a 360-degree wire-wound potentiometer with a motor-driven wiper arm and fixed voltages applied to both ends of the winding. In these units, the voltage at the wiper arm depends on the angular position of the shaft (input) to which the arm is attached.

The voltage changes abruptly as the arm moves from one wire to another, and also as the arm short-circuits adjacent wires. A typical input-output relation is shown in Fig. 4(A).

Initially, the arm can be just ready to assume a new potential (output) with a small change of shaft position, Fig. 4(B), can have just reached a new potential, Fig. 4(C), or can be in an intermediate position, Fig. 4(A).

The real part of the nondimensional describing function of a granular device depends on the initial condition; the reactive component is zero for all initial conditions. The linearized gain is considered a separate factor K_P. The describing functions for the various initial conditions are:

For Fig. (4A)

$$A_1=\frac{2}{n^2\pi}\left(\sum_{a=1}^{a=n}\sqrt{4n^2-(2a-1)^2}\right)$$

(for integer n's) (1)

For Fig. 4(B)

$$A_1=\frac{4}{n^2\pi}\left(\sum_{a=0}^{a=n-1}\sqrt{n^2-a^2}-\frac{n}{2}\right)$$

(for integer n's) (2)

Equation 2 is also the describing function for the condition shown in Fig. 4(C).

Fig. 5 plots the values of the describing

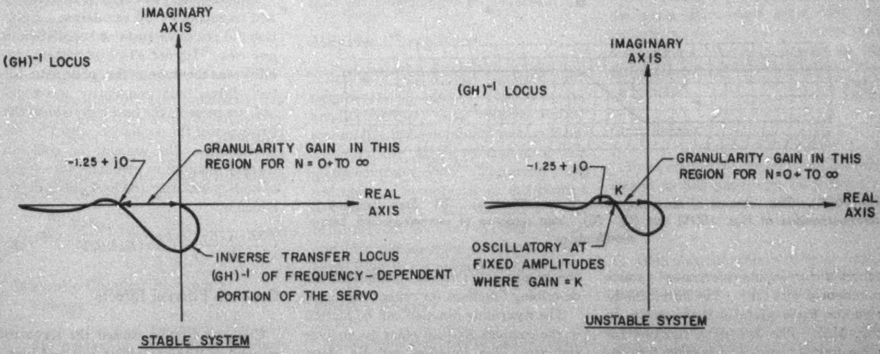

IMAGINARY AXIS
(GH)⁻¹ LOCUS
-1.25 + j0
GRANULARITY GAIN IN THIS REGION FOR N = 0+ TO ∞
REAL AXIS
INVERSE TRANSFER LOCUS (GH)⁻¹ OF FREQUENCY-DEPENDENT PORTION OF THE SERVO
STABLE SYSTEM
IMAGINARY AXIS
(GH)⁻¹ LOCUS
-1.25 + j0
K
GRANULARITY GAIN IN THIS REGION FOR N = 0+ TO ∞
REAL AXIS
OSCILLATORY AT FIXED AMPLITUDES WHERE GAIN = K
UNSTABLE SYSTEM

Fig. 7. Superposition of loci to show effect of granularity on a servomechanism's performance

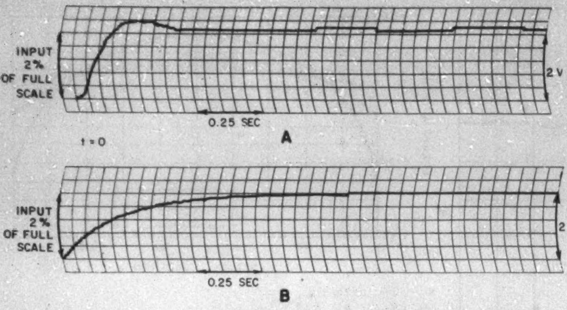


Fig. 8(A). Jitter caused by granularity. **(B).** Step response of system with granularity and sufficient coulomb friction

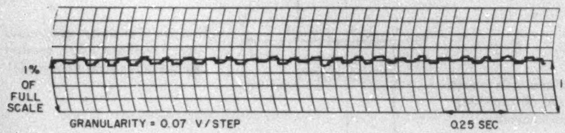


Fig. 9. Jitter caused by d-c unbalance

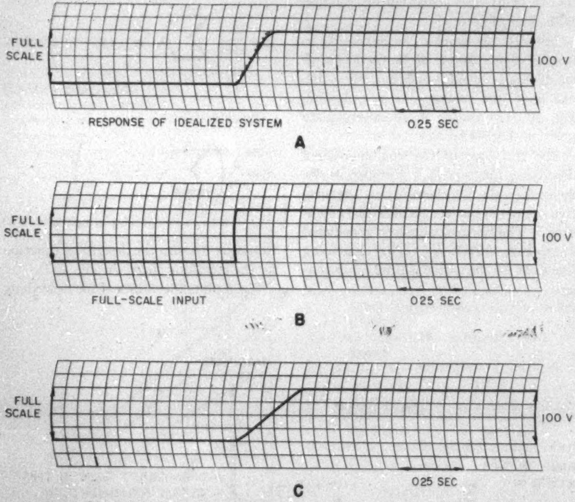


Fig. 10(A). Step response of servomechanism after modifications. **(B).** Step command to servomechanisms of Figs. 10(A) and (C). **(C).** Step response of servomechanism before modifications

function for inputs traversing various numbers of wire ($2n$). The initial condition for these plots is that shown in Fig. 4(A). The describing function for other starting conditions is given for constant n in Fig. 6. Interpolation between the graphs and formulas establishes the describing functions for other conditions.

The describing function can be plotted in the complex $KG(j\omega)$ plane to analyze stability. The function will lie along the negative real axis. The point(s) of intersection of the locus of the granularity-describing function and the locus of the frequency-dependent portion of the system will indicate the frequency of oscillations. A number of different amplitudes of oscillation are possible since the describing function is not everywhere unique. Fig. 7 illustrates the method of stability analysis. The loop gain, exclusive of the nondimensional describing function, can be adjusted to prevent intersection of the loci.

If granularity is present but coulomb friction and amplifier dead space are not the system may hunt between adjacent output levels if neither level corresponds to the input command. Fig. 8(A) shows the effect of jitter on the response to a step command. In Fig. 8(B), the granularity-caused jitter has been eliminated by addition of coulomb friction. The value of the friction necessary to accomplish this is shown by the simple relation

$$T_{SF} \geqq K_A K_T Q \tag{3A}$$

where K_T is defined by the stalled motor torque equation

$$T_A \cong K_T E_C \tag{3B}$$

Gear Backlash

Gear backlash can cause prolonged damped oscillation in the response of a servo to a step-command input when coulomb friction is negligible. The oscillation may resemble that caused by granularity. As coulomb friction is increased, the period of damped oscillation decreases. The amount of coulomb friction required to eliminate such damped oscillation is defined by

$$T_{SF} \cong \frac{BK_P K_A K_T}{N} \tag{4A}$$

Equation 4(A) was derived analytically and corroborated by simulation. Backlash did not cause sustained oscillation in any case. Its net effect on system stability was the same as that of an attenuator. When both granularity and backlash are present, the backlash reduces the frequency of the oscillation caused by the granularity. The amount of coulomb friction required to eliminate the effects of both granularity and backlash is determined by

$$\left(\frac{BK_P K_A K_T}{N}\right) < T_{SF} > (K_A K_T Q) \tag{4B}$$

Coulomb Friction Effects

Coulomb friction caused the servo to respond as if the amplifier gain had been lowered. If coulomb friction is increased

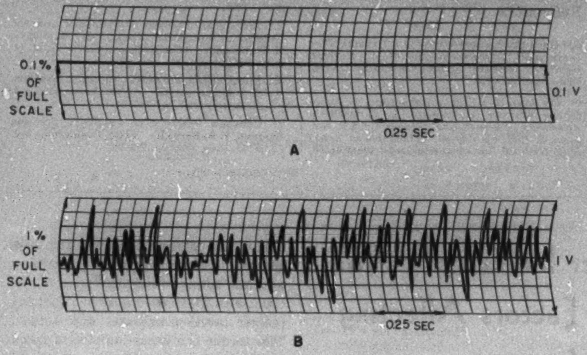


Fig. 11(A). Output noise in servomechanism after modifications. (B). Output noise in servomechanism before modifications

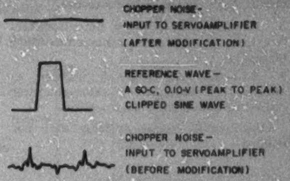


Fig. 12. Vibrator or chopper noise before and after modifications

to stop oscillations produced by granularity or backlash, or both, the response of the servo will be slowed and static error will be increased. A method of calculating the effect of coulomb friction on the speed of system response has been presented previously.[5]

D-C Static Unbalance Voltage

The d-c static unbalance voltage between the halves of the push-pull output stage of the amplifier combines with the 60-cycle reference voltage and produces a strong 60-cycle motor torque.

Although the 60-cycle torque was somewhat attenuated by the motor-load dynamics, the servo output contained a 60-cycle oscillation even in the presence of zero-input signal. The nonlinear characteristic of the motor produced a 30-cycle subharmonic in the output; see Fig. 9. The relationship for eliminating oscillation caused by d-c unbalance in the amplifier is

$$T_{SF} \cong \left| \frac{K_T E_{DC}}{1+j2\pi F_1 T_E} \right| \qquad (5A)$$

The presence of both granularity and amplifier unbalance requires more coulomb friction to prevent oscillation than does the presence of only one of these nonlinearities. This value of coulomb friction can be found by

$$T_{SF} \cong \left| \frac{K_T E_{DC}}{1+j2\pi F_1 T_E} \right| + K_A K_T Q \qquad (5B)$$

Mutually Coupled Reference Voltage

The portion of the 60-cycle reference winding voltage coupled by mutual inductance into the control winding causes a 120-cycle torque. Although the effect of the torque is greatly attenuated by the motor-load dynamics, the mutually inducted voltage will cause oscillation if the coulomb friction is barely sufficient to prevent oscillation caused by granularity and backlash.

Wiper-Arm Contact Noises

Several types of noise emanate from wire-wound potentiometers. One is caused by variations of contact resistance which introduces sharp spikes of voltage into the output and contributes to jitter.

Should the wiper arm traverse the potentiometer winding at a velocity in excess of the critical value, additional spikes will occur when the wiper arm leaves the winding because of centrifugal force. These spikes, in all cases observed, were severer than contact resistance noise, and their effects were more pronounced.

Amplifier Pickup Noise

Amplifier pickup noise (primarily line frequency) will cause a voltage in the amplifier output. This voltage is often appreciable in high-gain servo amplifiers. The amplifier output is applied to the motor control winding. The motor shaft will oscillate in response to this control winding voltage. The oscillation tends to dither out coulomb friction. If the amplifier pickup noise is severe, it will cause jitter in the output.

Vibrator Noise

Synchronous vibrator or chopper noises of frequencies determined by the chopper-reed driving frequency and mechanical resonance can cause important output noise if the impedance of the driving voltage source is high and the amplifier gain is large. Simulation investigation supported this contention.

Modification of a Commercial Instrument Servo

A production model of an instrument servo was modified on the basis of the principles outlined in the foregoing. In this servo, the jitter was caused principally by d-c static unbalance in the amplifier. To keep the jitter to a tolerable level, the amplifier gain had to be restricted below the design value. When the jitter was eliminated, it was possible to raise the amplifier gain, and also to linearize the gain for larger input swings.

Fig. 10 contrasts the servo response to a full-scale step input before and after modification. For comparison, the response of an idealized servo is also shown. The idealized servo was assumed to have the same velocity and acceleration limits as the actual servo. The acceleration was assumed to jump to the limit at the instant the input was applied; drop to zero at the instant the velocity limit was reached; drop instantly to a negative value equal to the acceleration limit as the error approached zero; and rise abruptly to zero at the instant the error reached zero. For large step inputs, the actual servo closely matched the idealized performance.

Conclusions

Jitter caused by the nonlinearities discussed in this paper can be reduced or eliminated, as shown in Figs. 11(A) and (B), by:

1. Obtaining the vibrator driving voltage from a low-impedance source such as a filament transformer; see Fig. 12.
2. Reducing the d-c static unbalance between the halves of the push-pull output stage of the amplifier by inserting a simple d-c degeneration network in the grid-biasing

circuit, or by an equivalent measure.

3. Keeping the feedback potentiometers free from dirt to reduce contact resistance noises.

4. Adjusting the coulomb friction, as indicated in the paper, to eliminate jitter caused by unavoidable nonlinearities.

Conditionally stable systems were not studied.

References

1. GEDA SIMULATION STUDY OF A CARRIER-TYPE INSTRUMENT SERVOMECHANISM, R. L. Hovious. *GER-5779*, Goodyear Aircraft Corporation, Akron, Ohio, April 21, 1954.

2. ON SOME NON-LINEAR PHENOMENA IN REGULATORY SYSTEMS (in Russian), L. C. Goldfarb. *Avtomatika i Telemechanika*, Moscow, U.S.S.R., vol. 8, no. 5, 1947, pp. 349–83. Translated in *Report 1691*, National Bureau of Standards, Washington, D.C., 1952.

3. SINUSOIDAL ANALYSIS OF FEEDBACK CONTROL SYSTEMS CONTAINING NONLINEAR ELEMENTS, E. C. Johnson. *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 169–81.

4. LIMITING IN FEEDBACK CONTROL SYSTEMS, Ralph J. Kochenburger. *AIEE Transactions*, vol. 72, pt. II, July 1953, pp. 180–94.

5. COULOMB FRICTION IN FEEDBACK CONTROL SYSTEMS, V. B. Haas, Jr. *AIEE Transactions*, vol. 72, pt. II, May 1953, pp. 119–26.

Aircraft Tachometer Indicator—An Analysis of Design Factors Affecting Starting Performance

L. T. AKELEY
MEMBER AIEE

FIFTY years of powered flight have brought changes to aircraft controls. A pilot cannot properly control the flight of today's powerful machines without the use of instruments. Many of the newer instruments and control devices embody electronic elements for sensitivity and speed of response. Some have servosystem operation for highest accuracy. One of the oldest instruments, the engine speed indicator,[1] still has an important place on the aircraft instrument panel. Without electronic components or servo elements the simple dragmagnet–eddy-current disk type of tachometer instrument continues to provide a highly accurate and dependable service to the pilot.

In general, the aircraft tachometer indicator, reading revolutions per minute (rpm) engine speed or per-cent rated rpm, serves as a watchdog on safety and economy. An engine usually has a recommended cruising rpm for maximum economy of fuel consumption; there are recommended rpm's for particular aircraft maneuvers or attitudes; there may be an rpm value which should not be exceeded to stay within the strength of rotating or reciprocating parts; and there is an rpm consistent with allowable temperature rise in the engine. These last two points have particular significance for the gas-turbine type of aircraft engine. The increased efficiency which accompanies higher speed and higher operating temperature has led to engine designs which have minimum margins of safety, placing greater importance and dependence upon the tachometer indicator.

Starting rpm is more critical on the gas-turbine engine than on the reciprocating engine. The danger of fire and explosion or other engine damage in starting is greater, particularly on the engine equipped with a slowly accelerating electric starter. Since "cartridge" and other combustion-type starters accelerate an engine four or five times as fast as the electric starter, the danger of flooding the engine and overtemperaturing the turbine is much less. Tachometer indication is therefore less important to the faster start. With the electric starter the aircraft's tachometer indicator must give the correct rpm readings needed by the pilot to follow a safe starting procedure. Obviously, with either starter the tachometer indicator must start and indicate properly somewhere below the minimum engine rpm reading that is required by the pilot.

General Discussion

Fig. 1 shows schematically the elements of a widely used aircraft tachometer system. In this system a 3-phase 2-pole voltage generator is mechanically coupled to the aircraft engine shaft through suitable gearing so that 100-per-cent rated speed of the aircraft engine corresponds to 4,200 rpm of the generator rotor. Three-phase power from the generator is carried electrically to the instrument panel where the indicator's 4-pole synchronous 3-phase motor is excited. A dragmagnet attached to the rotor shaft produces a torque proportional to speed of rotation on a lightweight eddy-current disk in the field of the rotating magnet. The disk-gear train-pointer system is restrained by a spiral spring which exerts a torque counter to the disk torque such that the pointer position indicates disk torque. This torque is a linear function of speed of the dragmagnet, which is an integral function of the generator frequency, which is a linear function of engine rpm. Therefore, pointer position may be used to indicate engine rpm or per-cent rated rpm.

Fig. 2 shows schematically the rotor construction used in the tachometer indicator motor. The rotor has two torque-producing elements, a permanent magnet J_m, and a hysteresis element J_h. In general, J_m provides starting and lock-in torque, whereas J_h provides higher frequency torque for accelerating the rotor close enough to synchronism for the permanent-magnet rotor to pull in. J_d is the aforementioned dragmagnet and shaft. Not shown are coupling pins between J_m, J_h, and J_d, which tie them together so that the dragmagnet J_d will be driven synchronous with J_m once starting is completed.

There are many design factors affecting starting performance of a synchronous motor. Concordia, Crary, Kilbourne, and Weygandt[2] have set up equations in a form suitable for solution on the differential analyzer; they present curves for both solution and confirming tests. However, since a tachometer indicator rotor differs from the usual synchronous rotor design and since tachometer starting conditions differ from their conditions, another approach to the problem seems to be warranted. Results of the following analysis provide a means for design calculations. Assumptions are made where indicated to simplify the solutions for design application.

The Nomenclature gives a list of

Paper 54-358, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 16, 1954; made available for printing August 6, 1954.

L. T. AKELEY is with the General Electric Company, Lynn, Mass.

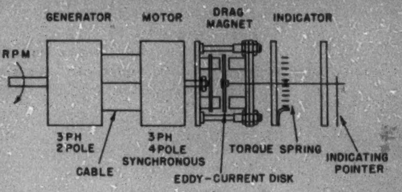


Fig. 1 (left). Schematic of aircraft tachometer system

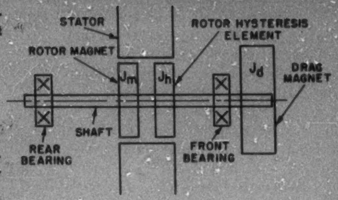


Fig. 2 (right). Schematic of tachometer indicator motor

motor design parameters, symbols, and other definitions, as well as a list of motor parameter values which will be used to demonstrate use of the derived equations. As previously indicated, Fig. 1 shows the tachometer system being considered. Fig. 2 shows the particular motor elements which are pertinent to the motor-starting analysis which follows.

Nomenclature

Symbols and Terminology

T_a = peak accelerating torque seen by the rotor
T_m = peak gross motor torque resulting from reaction of stator and rotor fields
T_L = peak locking torque of the rotor magnet to the stator
T_F = average friction torque of the rotor bearings
T_u = peak rotor mechanical unbalance torque
T_d = disk reaction torque seen by the rotor
$\mathfrak{T}$ = average or effective torque
a = average acceleration
A_s = stator area per pole = $(\pi DS)/2P$
D = diameter or motor bore
S = motor stack length
P = number of pairs of motor poles
t = time
α = angular displacement between stator and rotor fields
θ = general symbol for angle
E_m = rms volts per phase induced in the stator by rotation of the rotor magnet
f_m = frequency, in cycles per second (cps), used in determining E_m $= \dfrac{P \times \text{driven rotor rpm}}{60}$
E_g = minimum generator rms volts per rpm per phase with motor load
Z = motor impedance per phase
f = generator exciting frequency, cps
I = stator coil current
J = rotor inertia
J_m = inertia of the rotor magnet
J_h = inertia of the hysteresis disk
J_d = inertia of the dragmagnet
B_r = the peak flux density produced by the rotor magnet as seen by the stator coils
Δ_1 = stator rms ampere-turns per unit length of periphery
p = stator pitch
k_p = stator pitch factor = $\sin (P\pi)/2$
H_f = rotor hysteresis torque, a function of excitation as indicated in Fig. 3
k_d = a constant coefficient for eddy-current disk reaction torque
G = a motor torque coefficient defined in Appendix I

Assumed Motor Parameters for Starting Calculations

T_L = 0.65 gram-centimeter (gm-cm)
T_F = 0.20 gm-cm
T_u = 0.40 gm-cm
J_m = 0.004 gm-cm-second^2
J_h = 0.004 gm-cm-second^2
J_d = 0.050 gm-cm-second^2
G = 0.77 gm-cm per cycle per second (calculated from equation 22)
P = 2
E_g = $0.0035/\sqrt{3}$ volts per rpm per phase

Analysis of Motor Starting

Consider a motor having a rotor and shaft with dragmagnet inertia on the end of the shaft, two bearings supporting the shaft, and a stator, as indicated schematically in Fig. 2. If the stator is excited, an equation for torque equilibrium may be stated as follows

$$\mathfrak{T}_m + \mathfrak{T}_a + \mathfrak{T}_F + \mathfrak{T}_L + \mathfrak{T}_u + \mathfrak{T}_d = 0 \qquad (1)$$

where the above $\mathfrak{T}$ terms represent effective values of gross motor torque, net accelerating torque, friction torque, locking torque, unbalance torque, and disk reaction torque (in that order). A rigorous statement of torque equilibrium would require that equation 1 be a differential equation, not solvable by simple methods. The simplifying assumption that effective values of the various terms may be stated and added together as in equation 1 is reasonable only because some terms are much larger than others. In the motor to be considered not all of these terms act at the same time and they are not equally significant to motor starting. However, looking in detail at each of the terms of equation 1

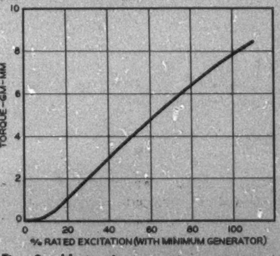


Fig. 3. Hysteresis torque versus per-cent excitation for a tachometer indicator motor

$$\mathfrak{T}_m = \frac{2G}{\pi} f + H_f \qquad (2)$$

where

f = motor excitation frequency, cps
G = a coefficient (see Appendix I) such that $G \sin P\alpha f$ is motor torque resulting from reaction of stator and permanent-magnet rotor fields
H_f = hysteresis torque produced in the hysteresis element of the rotor (described quantitatively in Fig. 3)

$$\mathfrak{T}_{a_L} = -3\pi J f^2/P \qquad (3)$$

is the average accelerating torque required to accelerate to synchronism a rotor restrained only by magnetic locking torque (see Appendix II for derivation), where

J = rotor inertia
P = number of pairs of poles in the motor
f = motor excitation frequency at which the rotor pulls into synchronism

$$\mathfrak{T}_{a_u} = -4\pi J f^2/3P \qquad (4)$$

is the average acceleration torque required to accelerate to synchronism from rest a rotor restrained only by rotor unbalance; see Appendix II. This differs from equation 3 because the angle of a complete locking torque cycle is smaller than that of unbalance torque.

$\mathfrak{T}_F = T_F$ is average rotor friction, assumed to be a constant for a particular bearing and bearing load.

$$\mathfrak{T}_L = -2T_L/\pi \qquad (5)$$

is the average magnetic locking torque to the stator of the permanent-magnet rotor where T_L is the peak value of a locking torque which is assumed sinusoidal; see Appendix II.

$$\mathfrak{T}_u = -2T_u/\pi \qquad (6)$$

is the average torque on the rotor due to mechanical unbalance, where T_u is peak unbalance torque of a sinusoidal term.

$$\mathfrak{T}_d = -K_d \frac{d\theta}{dt} \qquad (7)$$

is reaction torque from the eddy-current disk, proportional to the speed of the dragmagnet with respect to the disk. K_d is a constant.

Equations 2 through 7 can be substituted into equation 1 to describe all the torques that may be imposed upon the rotor. The rotor is so constructed, however, that not all of these restraining torques act at the same time. This fact plus some simplifying assumptions allows solving equation 1 as an algebraic equation.

Fig. 3 shows hysteresis torque H_f for an average rotor plotted against per-cent motor excitation. Since generator voltage output is proportional to rpm, the per-cent rated excitation represents both frequency and voltage. As indicated in Fig. 3, hysteresis torque at low frequencies, say below 2-per-cent rated speed, is very small and will be assumed zero for low-frequency starting calculations. H_f is a nonlinear function of excitation such that hysteresis torque becomes appreciable at higher frequencies.

Reaction torque T_d is easily calculated from design information on the instrument spring. At low frequencies it is very small and will be assumed zero for starting calculations. Locking, friction, and unbalance torques are significant and will be further evaluated.

The permanent-magnet part of the rotor has a locking torque T_L to the stator which varies with magnet strength and with position of the rotor-magnet poles with respect to the stator slots. Locking torque exists because the stator slots produce a variation in reluctance of the flux path of the rotor-magnetic field with change in angular position of the rotor magnet. Rotor-stator configuration and materials can effect considerable difference in locking torque. Peak locking torque T_L can be measured, the number of peaks per rotor revolution being equal to the number of stator slots. Peak mechanical unbalance torque T_u is easily measured; mechanical unbalance is a sinusoidal function of rotor position. Both unbalance and locking torques have negative as well as positive peaks—energy absorbed in one half-cycle is returned in the next half-cycle.

Rewriting equation 1, assuming that hysteresis and reaction torques are zero and substituting equations 2, 3, 5, and 6 into equation 1

$$\frac{2G}{\pi}f-\frac{3\pi J}{P}f^2-T_F-\frac{2T_L}{\pi}-\frac{2T_u}{\pi}=0 \tag{8}$$

In writing equation 8, a binomial equation, the simplifying assumption is made that effective torques can be added algebraically. Actually locking and unbalance torques are sinusoidal functions of rotor position. However, since locking torque peaks positive 12 times per revolution while unbalance is peaking once, it is apparent that as the rotor approaches its peak unbalance position it will have at least cos $\pi/12$ times its peak value coinciding with peak locking torque. Thus, it would be more exact in some cases to add peak unbalance instead of average unbalance torque to average friction and locking torques. However, since unbalance is usually small when the very significant locking torque is being considered and since, where unbalance is significant, average unbalance torque is pertinent with the better rotor constructions to be analyzed, for simplicity of application average unbalance torque is used in equation 8. Likewise, since locking torque governs the starting of a motor having low rotor unbalance, accelerating torque in a form consistent with locking conditions (see Appendix II) is used in equation 8.

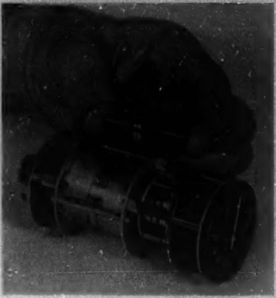


Fig. 4. Type DJ43 tachometer indicator, showing the rotor with permanent-magnet and hysteresis elements free on the shaft

Equation 8 has the form

$$A'f^2+B'f+C'=0 \tag{9}$$

where

$A'=-3\pi J/P$ for the case where locking torque is paramount (10)

$$B'=2G/\pi \tag{11}$$

$$C'=-T_F-2T_L/\pi-2T_u/\pi \tag{12}$$

Equation 9 may be solved for motor starting frequency

$$f=\frac{-B'+\sqrt{B'^2-4A'C'}}{2A'} \tag{13}$$

A', B', and C' are calculable or measurable and are subject to design variation for the purpose of achieving better motor starting performance. Substitution of parameters into equation 13 such that the solution of f is imaginary indicates that the motor will not start under the conditions assumed.

From equation 13 $f_{\max}$, the maximum frequency at which a rotor will start synchronously as defined in equation 3 and Appendix II, is

$$f_{\max}=-B'/2A' \tag{14}$$

for which case

$$C_{\max}'=B'^2/4A' \tag{15}$$

The fact that there is an upper frequency of starting, $f_{\max}$, as well as a minimum starting frequency, equation 13, is borne out by observation. For example, a motor without hysteresis disk will start and run synchronously at some minimum value of excitation but fail to start above some higher value corresponding to $f_{\max}$. Equation 15 may be considered a "figure of merit" for a motor design, since it tells how large the restraining torques (locking, unbalance, and friction) may be and still permit starting.

Referring to Fig. 2, there are three significant inertia elements in the rotor: the rotor magnet J_m, the hysteresis element J_h, and the dragmagnet J_d. Inertia of the shaft is small and will be ignored in design calculations. J_m and J_h have about the same inertia; both are much smaller than J_d. Since J_d is to be driven synchronously by J_m and J_h, the three elements must be mechanically attached by some means to each other or to the common shaft. Suppose, first, that all three are attached solidly to the shaft. Such a rotor would not start well, since all three inertia elements must be accelerated at the same time. A better construction, second, attaches J_h and J_d solidly to the shaft and permits J_m slightly less than 360 degrees' freedom on the shaft. This construction has been used in tachometer indicators for a long time. A still better construction, third, found in a current design, has both J_m and J_d free on the shaft such that the rotor starts as follows. J_m, which derives its torque from the reaction of its own field with the stator ampere-turns, first accelerates to synchronism before it is coupled mechanically to J_h by means of interfering pins on the two elements (see Fig. 4); J_h is then accelerated to synchronism before it is coupled to the shaft and to J_d, which is solidly attached to the shaft; J_d is then accelerated to synchronism. These actions will occur in series in the course of normal starting.

Sample motors built as described here have been tested and the nature of their starting observed. Observations confirm, e.g., that the starting actions of the third construction occur in series, as just

Table I. Summary of Data for Three Constructions

Construction	Assumed Values of Inertia and Load							Calculated					
	J_m	J_h	J_d	J	T_F	T_L	T_u	A'	B'	C'	Starting f, cps	C_{max}'	f_{max}, cps
First: Assuming that All Three Rotor Elements are Attached Solidly to Shaft													
	0.004	0.004	0.050	0.058	0.2	0.65	0.4	−0.273	0.490	−0.84	Imaginary Will not start.	−0.220	0.90
Second: Assuming that Rotor Magnet J_m is Free on Shaft to Start First													
J_m alone starts	0.004			0.004	0.05	0.65	0.05	−0.01885	0.490	−0.495	1.05	−3.2	13.0
J_h and J_d start		0.004	0.050	0.054	0.20	0	0.40	−0.1131	0.490	−0.454	1.34	−0.53	2.2
Third: Assuming that Both J_m and J_h are Free on Shaft to Start One at a Time													
J_m alone starts	0.004			0.004	0.05	0.65	0.05	−0.01885	0.490	−0.495	1.05	−3.2	13.0
J_h starts next		0.004		0.004	0.10	0	0.10	−0.006283	0.490	−0.164	0.32	−9.6	39
J_d finally starts			0.050	0.050	0.20	0	0.40	−0.1047	0.490	−0.454	1.27	−0.57	2.34
Assuming Zero Friction and Unbalance for Third Construction													
J_m alone starts	0.004			0.004	0	0.65	0	−0.01885	0.490	−0.4138	0.87		
J_h starts next		0.004		0.004	0	0	0				0		
J_d finally starts			0.050	0.050	0	0	0				0		

stated. Unbalance and friction of the dragmagnet, however, may be such that the rotor magnet will be pulled out of synchronism, reverse its direction, start again, and possibly repeat the starting sequence several times before getting the dragmagnet up to synchronous speed.

Locking torque may account for more than 50 per cent of C', a fact which explains the great advantage of having J_m free on the shaft to start first and overcome the locking torque before meeting the major inertia and unbalance of J_d. A ball bearing is used at J_m to assure low friction to the shaft.

Using assumed and calculated design values as given in the Nomenclature, equation 13 is used to calculate starting frequency for the three conditions just described. Table I summarizes these calculations for the three constructions. All inertia terms have been calculated or measured. Peak locking torque is a measured value for an average rotor. Friction and unbalance torque values assumed are based on measurements on a few samples; they are more likely to vary from sample to sample than is the locking torque. The value of G, the motor torque coefficient, was calculated by substituting known motor and generator parameters into equation 21. Values for all these parameters are not listed in the Nomenclature but are completely defined there.

Looking at Table I, for the first assumption that all three rotor elements are solidly attached to the shaft, the solution for f (equation 13) is imaginary, indicating that the rotor will not start for the assumed conditions. Calculated f_{max} and C_{max}' from equations 14 and 15 indicate that the rotor would start at 0.90 cps, or 54 cycles per minute (cpm), if C' had a value of −0.220. This is impossible with the high locking torque which exists.

For the second assumption that the rotor magnet J_m is free to move approximately 180 degrees on the shaft before coupling mechanically to the hysteresis element and the dragmagnet fixed to the shaft (a ball bearing is used on the rotor magnet to assure low friction), calculated f is 1.05 cps (63 cpm) for the rotor magnet, which starts first, and 1.34 cps (80 cpm) for the rest of the rotor. In calculating this latter figure equation 4 is pertinent instead of equation 3 for accelerating torque so that $A' = -4\pi J/3P$ is used to calculate A' instead of equation 10. This is done whenever unbalance torque determines starting conditions instead of locking torque. Calculation of C_{max}' and f_{max} shows that the second construction has a much greater safety margin for starting than the first.

For the third assumption that both rotor magnet and hysteresis element are free on the shaft except for coupling pins between the three members (see Fig. 2), calculated starting frequency for the rotor magnet is 1.05 cps (63 cpm). With the rotor magnet started and locking torque overcome, the hysteresis element is easily accelerated to the synchronous speed of the rotor magnet. In fact, friction between hysteresis element and the shaft can be reasonably high (no ball bearing required) because locking torque is zero and inertia of J_h is low. The dragmagnet and shaft J_d is next subjected to the starting torque and, as indicated in Table I, calculated starting frequency for J_d is 1.27 cps (76 cpm). Calculated C_{max}' and f_{max} indicates a greater safety factor for the third construction than for the second.

The last section of Table I shows calculated starting frequency for the third construction, assuming zero unbalance of all parts and zero friction for all bearings. For these ideal conditions locking torque determines the starting frequency of the rotor, calculated to be 0.87 cps (52 cpm). One would not expect a motor to start below this figure.

Test Results

Samples have been built and tested with all three constructions described in the foregoing. Visual observation of motor starting with different rotor constructions and with different values of dragmagnet inertia, locking, unbalance, and friction

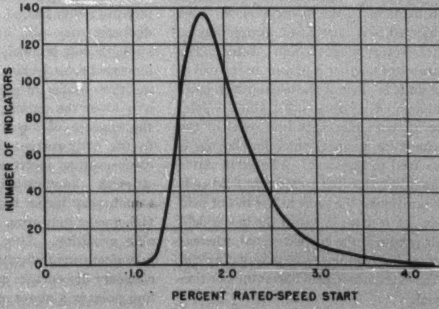


Fig. 5. Low-frequency starting tests on 581 type DJ43 tachometer indicators (production record)

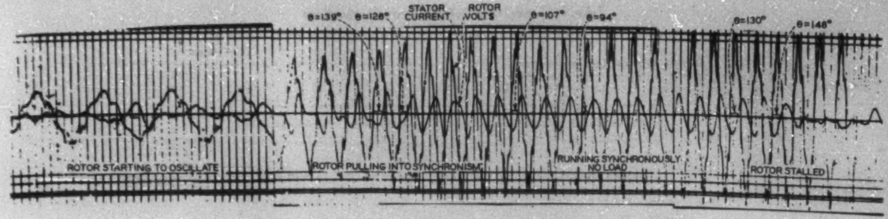


Fig. 6. Oscillographic record of a tachometer motor starting, running, and stalling

torques has provided experimental evidence of the relative importance of these items to low-frequency starting. These tests also show that the third construction is superior to the second, and the second, superior to the first.

581 production tachometer indicators similar to the one shown in Fig. 4, having both rotor magnet and hysteresis elements free on the shaft, were tested for low-frequency starting. Fig. 5 is a plot of these test results. This particular factory test is made at room temperature with a generator output (with load) of $0.0035/\sqrt{3}$ volts per phase per rpm. All of the calculated starting frequencies were based upon a similar generator output.

Fig. 6 is an oscillographic record of a tachometer indicator motor starting, running no load, and stalling at low-frequency excitation. One trace records stator current in a single phase; the other trace records voltage induced in a search coil from the rotor magnet field only. Assuming that the stator field flux is in phase with the current and that induced voltage from the rotor magnet field is proportional to rate of change of field flux, it follows that the angle between the two fields is the measured angle minus 90 degrees. Note that this angle, $\alpha=\theta-90$ degrees, is approximately 45 degrees at start and stall, whereas, as expected, it is close to zero running no load. This test was made using the motor which is to be subjected to a starting excitation plus an extra rotor magnet and stator. The extra rotor magnet is placed on the end of the shaft in place of the usual dragmagnet and angularly positioned so that its poles line up with corresponding poles of the motor rotor magnet, which for this test is also fixed to the shaft. The extra stator is angularly located so that its field coils are positioned the same as the motor field coils. A resistor in one of the motor field coils produces the current signal, whereas voltage induced in the corresponding coil of the extra stator provides rotor-position signal.

General Conclusions

These calculations plus experimental tests indicate that the third construction with both the permanent magnet and hysteresis rotor elements free on the shaft (the construction shown in Fig. 4) is best, being far superior in low-frequency starting to the first construction and significantly better than the second construction, which has only the permanent magnet free on the shaft.

Fig. 5 test results show a mode starting frequency of about 1.75-per-cent rated speed, or 74 cpm motor excitation frequency. The minimum observed starting frequency was 1-per-cent rated speed, 42 cpm. These test results show reasonable agreement with calculated starting frequencies. The large spread of Fig. 5 indicates considerable variation between units, probably mostly due to differences in their unbalance and friction torques.

Equations 1 through 15 may be analyzed to show what may be done to further improve motor starting performance. Lower rotor inertia and higher stator ampere-turns will definitely lower starting frequency. Proper motor and generator impedance match will provide maximum ampere-turns in the motor. Limits on allowable current and heating at maximum excitation, however, do limit the ampere-turns in a given motor size. Increase in rotor diameter and stack should increase motor torque. Increase in rotor diameter may also increase rotor field strength, but probably at the expense of increased rotor locking torque. Increasing the number of poles would presumably lower the starting frequency, but at the expense of lower torque delivered to the eddy-current disk, assuming no compensating increase in dragmagnet strength. Sacrifice of torque at this point would mean higher friction errors in the tachometer indicating head, an undesirable condition. Also, probable greater manufacturing difficulty and lower motor efficiency discourages design of more than four poles in a motor of this size.

Appendix I. Derivation of Effective Gross Motor Torque $\mathfrak{T}_m$, Equation 2

Consider the torque[3] resulting from the interaction of the stator ampere-turns and the field of the rotor permanent-magnet element only

$$\text{Torque}=\frac{\pi\sqrt{2}D^2SB_r\Delta_1\sin P\alpha}{40}\quad\text{dyne-centimeters}\qquad(16)$$

where

$$B_r=E_m\times10^8/4.44Nf_mA_s\left(\frac{2}{\pi}\right)k_p\quad\text{lines per square centimeter}\qquad(17)$$

$$\Delta_1=2NIK_p\times\text{number of phases}/\pi D\quad\text{ampere-turns per centimeter}\qquad(18)$$

where k_p = stator pitch factor = $\sin\pi p/2$, where p = stator pitch. Motor current per phase I can be stated in terms of the motor impedance per phase and the generator exciting voltage–speed characteristics as follows

$$I=E_g60f/Z\text{ amperes}\qquad(19)$$

$$A_s=\pi DS/2P\text{ square centimeters}\qquad(20)$$

Substituting equation 20 into 17, 19 into 18, and 17 and 18 into equation 16 and changing force units to grams, gives the following equation for gross motor torque in terms of motor parameters, speed-voltage characteristic of the generator with load, and the motor excitation frequency

$$\text{Torque}=\frac{0.974\times10^5PE_mE_g\text{ phases }\sin P\alpha\ f\text{ gm-cm}}{f_mZ}\qquad(21)$$

Adding to this torque, due to the permanent-magnet element of the rotor, the torque H_f from the hysteresis rotor element (as described in Fig. 3) gives an equation for gross motor torque for any frequency of excitation

$$T_m=G\sin P\alpha f+H_f\qquad(22)$$

where G is the coefficient of $\sin(P\alpha)f$ as defined in equation 21.

Assuming that hysteresis torque is much smaller than permanent-magnet rotor torque (true at low frequencies), average or effective motor torque

$$\mathfrak{T}_m=\frac{2G}{\pi}f+H_f \qquad (23)$$

$2G/\pi$ as an effective value of $G \sin P\alpha$ is derived as in equation 25. This approximation is true only for minimum starting conditions where field torque acts positively upon the rotor over a complete 180-degree range of α.

Appendix II. Consideration of Permanent-Magnet Rotor Starting

The rotating field of a 3-phase stator has an angular velocity

$$W_f=2\pi f/P \qquad (24)$$

Assume that rotor magnet unbalance is negligible compared to locking torque. This is usually true since rotors can be mechanically balanced, whereas there is no simple way to appreciably reduce locking torque without significant reduction in motor efficiency.

Consider the work done to overcome locking torque. Actually the rotor oscillates with the frequency of the rotating field till it reaches a speed and angle of oscillation to go "over the hump" and reverse the sign of locking torque. The peak kinetic energy of oscillation is equal, however, to the potential energy which would be acquired by the rotor, if the rotor were moved through one half-cycle of its locking characteristics. Locking torque being a function of stator coil slots peaks every 30 mechanical degrees (there are 12 stator slots). Assuming that locking torque is sinusoidal, the work done in moving through a half-cycle is

$$\mathfrak{T}_L\pi/12=\int_0^{\pi/12} T_L \sin 12\theta d\theta=-2T_L/12 \qquad (25)$$

where

T_L = peak locking torque

$\mathfrak{T}_L$ = average or effective locking torque over the half-cycle

From equation 25

$$\mathfrak{T}_L=-2T_L/\pi \qquad (26)$$

In order for the rotor to reach synchronism by definition, it must reach an angular velocity w_r equal to w_f. It must accelerate to w_f in a time such that the torque seen by the rotor will not reverse sign. Motor torque is positive over 180 electrical degrees, 90 mechanical degrees, assuming the rotor were standing still. Since the rotor is oscillating through 30 mechanical degrees just before start, it sees positive torque over 90+30 mechanical degrees or over a time

$$t=\frac{90+30}{180}\,\frac{1}{f}=2/3f \text{ seconds} \qquad (27)$$

Assuming that the rotor starts from rest, which it effectively does even though oscillating

$$\mathfrak{a}=\frac{W_r}{t}=\frac{W_f}{t}=\frac{2\pi f}{Pt} \qquad (28)$$

Substituting equation 27 into 28

$$\mathfrak{a}=3\pi f^2/P \qquad (29)$$

Since torque equals inertia times acceleration

$$\mathfrak{T}_{a_L}=-3\pi Jf^2/P \qquad (30)$$

an expression for the average torque required to accelerate a rotor having inertia J, restrained by locking torque only, to accelerate to synchronism with the field.

Consider an inertia J restrained only by an unbalance torque $T_u \sin \theta$. As argued for locking torque in equations 25 and 26

$$\mathfrak{T}_u=-2T_u/\pi \qquad (31)$$

Unbalance torque peaks positively only once a revolution. Following the same reasoning used in the foregoing to obtain acceleration required to overcome locking torque, if the rotor is assumed at rest with the heavy side at the bottom and is to be accelerated to synchronous speed just as the unbalance torque changes sign, the time that motor torque remains positive is

$$t=\frac{90+180}{180}\,\frac{1}{f}=3/2f \text{ seconds} \qquad (32)$$

$$\mathfrak{T}_{a_u}=-4\pi Jf^2/3P \qquad (33)$$

is an expression for average torque required to accelerate a rotor, restrained only by unbalance torque and starting from rest, to accelerate to synchronism.

References

1. The Magnetic-Drag Tachometer, R. G. Ballard. *AIEE Transactions*, vol. 61, 1942, pp. 366–68.

2. Synchronous Starting of Generator and Motor, S. Concordia, S. B. Crary, C. E. Kilbourne, C. N. Weygandt, Jr. *AIEE Transactions (Electrical Engineering)*, vol. 64, Sept. 1945, pp. 629–34.

3. The Nature of Polyphase Induction Machines (book), P. L. Alger. John Wiley and Sons, Inc., New York, N. Y., 1951, p. 60.

4. Military Specification *MIL-2-7069*, Washington, D.C., Dec. 29, 1950.

Additions to z-Transformation Theory for Sampled-Data Systems

GLADWYN V. LAGO
MEMBER AIEE

THERE is one important class of feedback control systems which until recently has received but limited attention. The systems in this class are known as sampled-data systems. The distinguishing feature of this class of systems is that at one or more points in the system, signal information has the form of a train of pulses.

One of the most powerful mathematical methods for analysis of sampled-data systems is known as the *z*-transformation method. Ragazzini and Zadeh[1] give an excellent summary of this method. The present *z*-transformation theory is in its infancy and, as additions to this theory are developed, it will become an even more powerful tool for analysis and design of sampled-data systems. The purpose of this paper is to present a number of additions to the present *z*-transformation theory in the hope that they will add to the usefulness of the *z*-transformation method.

Because this paper presents a number of related topics rather than a single development, each of these topics will be discussed briefly before a more detailed analysis is presented. After a brief review of the *z*-transformation theory, a situation is discussed wherein the present *z*-transformation theory can lead to erroneous results. By use of the method described in the section entitled "Suggested Notation," this situation can be avoided.

The final and initial value theorems are obvious additions to *z*-transformation theory and are presented for use in later developments.

One of the major criticisms of *z*-transformation theory is that the behavior of the system between sampling instants is unknown. This uncertainty can be removed by a number of rather tedious devices. When these devices are used, the basic simplicity of the *z*-transformation method is lost. This paper outlines a

---

Paper 54-523, recommended by the AIEE Feedback Control Systems Committee, and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 11, 1954; made available for printing September 24, 1954.

Gladwyn V. Lago is with the University of Missouri, Columbia, Mo.

The work described in this paper was done at Purdue University while the author was on a Radio Corporation of America predoctoral fellowship in electronics as administered by the National Research Council. The author gratefully acknowledges the help received from his advisor, Prof. John G. Truxal, during the preparation of the thesis on which this paper is based.

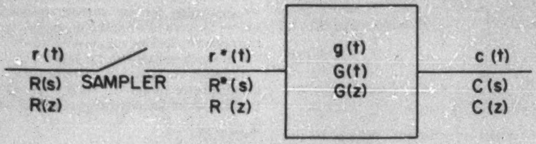


Fig. 1. Sampler followed by a network

method whereby the output at any point in the sampling period can be found by using a technique which preserves this basic simplicity.

This paper also outlines a technique for finding the pole-zero configuration of a continuous system coincides with a sampled-date system at the sampling instants. The current literature contains much material concerning the step-function response of continuous systems having various pole-zero configurations. It would be possible to duplicate this material for sampled-data systems in terms of z-transformation theory. However, the technique presented here eliminates this need and permits the designer already familiar with continuous systems to use his knowledge in determining the response of a sampled-data system.

The paper concludes by developing the equations for error coefficients in terms of the z-transforms. These equations play the same role in sampled-data theory as the conventional equations for error coefficients do in continuous systems.

Brief Review of z-Transformation Theory

The sampler is the component which distinguishes a sampled-data system from a continuous system. The input signal $r(t)$ to the sampler of Fig. 1 is continuous, whereas the output $r^*(t)$ is a train of pulses with a width T_1 seconds and a height proportional to the amplitude of $r(t)$ at the sampling instants. It is convenient mathematically to approximate this train of pulses by a corresponding train of impulses weighted so that the area of an impulse equals the pulse amplitude.[1] $R^*(s)$, the Laplace transform of $r^*(t)$, is given by two equivalent forms[2]

$$R^*(s)=\sum_{n=0}^{\infty} r(nt)\epsilon^{-nTs} \qquad (1)$$

and

$$R^*(s)=\frac{1}{T}\sum_{n=-\infty}^{+\infty} R(s+jn\omega_s)+\frac{r(0)}{2} \qquad (2)$$

where

$R(s)$ = the Laplace transform of $r(t)$
T = sampling period
$\omega_s = 2\pi/T$

$C^*(s)$, the Laplace transform of the output $c(t)$ of the network at the sampling instants, can be expressed as

$$C^*(s)=R^*(s)G^*(s) = \frac{1}{T}\sum_{n=-\infty}^{+\infty} C(s+jn\omega_s)+\frac{c(0)}{2} \qquad (3)$$

where $G^*(s)$ is defined as

$$G^*(s)=\frac{1}{T}\sum_{n=-\infty}^{+\infty} G(s+jn\omega_s)+\frac{g(0)}{2} \qquad (4)$$

Here, $C(s)$ and $G(s)$ are the Laplace transforms of $c(t)$ and $g(t)$ respectively .

$R(z)$, the z-transform of $r(t)$, is obtained by substituting z for ϵ^{sT} in equation 1 as

$$R(z)=\sum_{n=0}^{\infty} r(nT)\frac{1}{z^n} \qquad (5)$$

The z-transform of $r(t)=\epsilon^{-at}$ is taken to demonstrate the technique. The equivalent form to equation 1 is

$$R^*(s)=1+\epsilon^{-aT}\epsilon^{-sT}+\epsilon^{-2aT}\epsilon^{-2sT}+\ldots \qquad (6)$$

which can be written in closed form as

$$R^*(s)=\frac{\epsilon^{sT}}{\epsilon^{sT}-\epsilon^{-aT}} \qquad (7)$$

With the change in variables $z=\epsilon^{sT}$, equation 7 becomes

$$R(z)=\frac{z}{z-\epsilon^{-aT}} \qquad (8)$$

A short list of z-transforms is given in Table I.

Suggested Notation

The usual interpretation placed upon equation 4 is that $G(z)$ is taken from $g(t)$ in exactly the same manner as $R(z)$ is taken from $r(t)$. This means that $g(t)$ is first represented by a train of impulses at the sampling instants and then the usual steps are performed on the train of impulses to obtain the corresponding z-transform. As can be seen by the example in Appendix, this interpretation breaks down when $g(t)$ contains an impulse. This difficulty arises because the mathematics cannot distinguish between the impulse used to represent an amplitude and the true impulse.

Another interpretation of equation 4 is possible. The signal $r(t)$ becomes a train of impulses $r^*(t)$ as a result of the action of the sampler and not as a result of the fact that the z-transformation is taken. Since $g(t)$ is continuous (except for a possible impulse at $t=0$) and is not sampled, it does not follow that it must be converted to a train of impulses before taking the z-transform. One possible solution would be to let the points at the sampling instants, Fig. 2(B), be used to represent the $g(t)$ curve; see Fig. 2(A). However, since it is impossible to take the Laplace transform of a set of points, this cannot be done. If each point is held for a short

Fig. 2. Graphical depiction of g(t) and how g(t) can be represented by a train of points or pulses

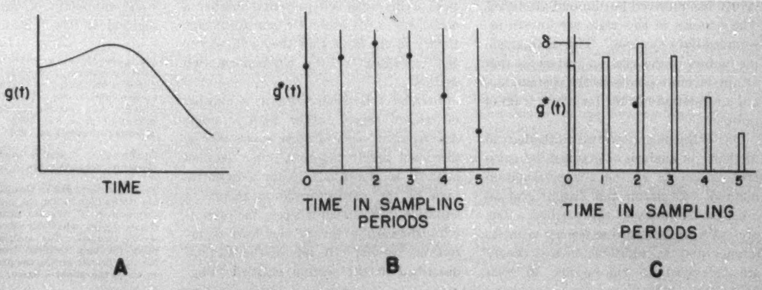


Table I. Short List of z-Transforms

Time Function	Laplace Transform	z-Transform
$u_0(t)$	1	1
$u_{-1}(t)$	$\frac{1}{s}$	$\frac{z}{z-1}$
t	$\frac{1}{s^2}$	$\frac{Tz}{(z-1)^2}$
ϵ^{-at}	$\frac{1}{s+a}$	$\frac{z}{z-\epsilon^{-aT}}$

time, δ seconds, Fig. 2(C), the Laplace transformation can be taken. A new table of z-transforms can be derived based upon this interpretation which is demonstrated with the example $g(t)=\epsilon^{-at}$. $G^*(s)$ can be written as

$$G^*(s)=1\left(\frac{1-\epsilon^{-\delta s}}{s}\right)+\epsilon^{-aT}\epsilon^{-sT}\left(\frac{1-\epsilon^{-\delta s}}{s}\right)+\epsilon^{-2aT}\epsilon^{-2sT}\left(\frac{1-\epsilon^{-\delta s}}{s}\right)+\ldots \tag{9}$$

The term in the parentheses can be factored out, the infinite series term can be summed, and with the change of variable $z=\epsilon^{sT}$ equation 9 can be written as

$$G(z)=\frac{z}{z-\epsilon^{-aT}}V \tag{10}$$

where

$$\left.\frac{1-\epsilon^{-\delta s}}{s}\right|_{s=\frac{1}{T}\ln z}=V \tag{11}$$

The symbol V can be thought of as an operator signifying that $G(z)$ represents values instead of impulses. It can be seen that the new table of z-transforms is exactly the same as Table I with the exception that each term is multiplied by V. The Appendix contains an example demonstrating the use of this notation. The only situation where this notation needs to be used is that in which the numerator and denominator of $G(s)$ are of the same degree in s. In the remainder of this paper it is assumed that the numerator of $G(s)$ is of lower degree than the denominator.

Final and Initial Value Theorems

Any z-transform of the output of a system $C(z)$ can be expanded in partial fractions as

$$C(z)=\frac{A_0z}{z-1}+\frac{A_1z}{z-\epsilon^{-a_1T}}+\frac{A_2z}{z-\epsilon^{-a_2T}}+\ldots \tag{12}$$

The inverse transform into the time domain is

$$c(t)=A_0+A_1\epsilon^{-a_1T}+A_2\epsilon^{-a_2T}+\ldots \tag{13}$$

If the Laplace transform of equation 13 has no poles in the right half of the s-plane or on the $j\omega$ axis, all terms of equation 13 go to zero as t goes to infinity with the exception of the A_0 term. Therefore, under these conditions (these same conditions apply to the final value theorem in the Laplace transformation theory) the final value of $c(t)$ can be found directly from $C(z)$ as

$$\lim_{z\to 1}\frac{(z-1)}{z}C(z)=\lim_{t\to\infty}c(t) \tag{14}$$

The initial value theorem can be derived by considering that $C(z)$ can be expanded in an infinite series in descending powers of z as

$$C(z)=B_0+\frac{B_1}{z}+\frac{B_2}{z^2}+\ldots \tag{15}$$

where

B_0 = value of $c(t)$ at $t=0$
B_1 = value of $c(t)$ at $t=T$, etc.

B_0 can be found as

$$\lim_{z\to\infty}C(z)=\lim_{t\to 0}c(t) \tag{16}$$

If the initial value of $c(t)$ is zero, an extension of the theorem is possible to give the value of $c(t)$ at $t=T$ in the form

$$\lim_{z\to\infty}(z)C(z)=c(t)\text{ at }t=T \tag{17}$$

The Output at Any Point in the Sampling Period

In the past, z-transformation theory has suffered from the uncertainty surrounding the behavior of the system between sampling instants. This uncertainty is partially removed by the procedure of finding the output at a submultiple of the sampling period.[8] The technique to be presented can be used to find the output at any time, Tx seconds, after the sampling instants. The system to be considered is that of Fig. 3(A) while Fig. 3(B) shows the hypothetical circuit after the artifices to be used in this technique are added. The signal $e^*(t)$ is a train of impulses at the sampling instants. $E(z)$ contains all the information that is contained in $e^*(t)$. However, all information except that at the sampling instants is discarded when $G(z)$ is obtained from $g(t)$. When $E(z)$ and $G(z)$ are multiplied together in the z domain, the corresponding occurrence in the time domain is the convolution of the train of impulses of $E(z)$ with the train of impulses (or values) of $G(z)$. Consequently, the same type of operation can be performed by obtaining a new function $G_{Tx}(z)$ from values of $g(t)$ at a set of points that occur T_x seconds after the sampling instants. The example $g(t)=\epsilon^{-aT}$ is used to demonstrate this method. The infinite series form for $G_{Tx}{}^*s$ is

$$G_{Tx}{}^*s=\epsilon^{-aTx}+\epsilon^{-a(T+Tx)}\epsilon^{-sT}+\epsilon^{-a(2T+Tx)}\epsilon^{-2sT}+\ldots \tag{18}$$

When the ϵ^{-aTx} term is factored out, the infinite series term can be summed and, with the change of variable $z=\epsilon^{sT}$, equation 18 can be written as

Table II. Short List of Modified z-Transforms

$G(z)$	$G_{Tx}(z)$
$\frac{z}{z-1}$	$\frac{z}{z-1}$
$\frac{Tz}{(z-1)^2}$	$\frac{Txz}{z-1}+\frac{Tz}{(z-1)^2}$
$\frac{z}{z-\epsilon^{-aT}}$	$\frac{\epsilon^{-aTx}z}{z-\epsilon^{-aT}}$

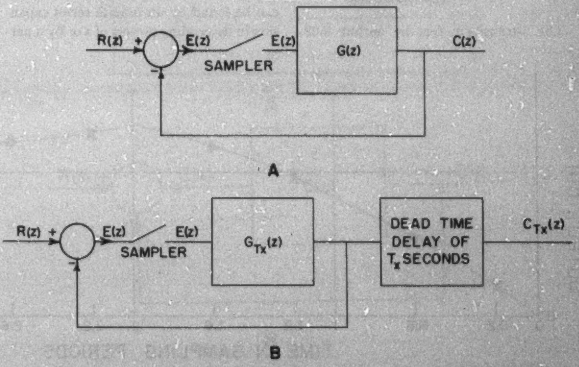


Fig. 3. Sampled-data system and hypothetical system used to find output at any time

$$G_{Tx}(z)=\frac{\epsilon^{-aTx}z}{z-\epsilon^{-aT}} \qquad (19)$$

A short list of z-transforms and the corresponding modified z-transforms are given in Table II.

The example to be considered to demonstrate this technique is the circuit of Fig. 3(B) with

$$G(s)=\frac{13.4}{s(s+10)}=1.34\left(\frac{1}{s}-\frac{1}{s+10}\right); \quad T=0.1 \text{ second} \qquad (20)$$

$G(z)$ is found as

$$G(z)=1.34\left(\frac{z}{z-1}-\frac{z}{z-\epsilon^{-1}}\right)=\frac{0.846z}{z^2-1.368z+0.368} \qquad (21)$$

When the input $r(t)$ is a unit step function, $E(z)$ is found as

$$E(z)=\frac{R(z)}{1+G(z)}=\frac{z^2-0.368z}{z^2-0.522z+0.368} \qquad (22)$$

The usual z-transform for the output is given by

$$C(z)=E(z)G(z)=\frac{z}{(z-1)}\,\frac{0.846z}{(z^2-0.522z+0.368)} \qquad (23)$$

The first few terms of the solution are placed in evidence by dividing the denominator into the numerator as

$$C(z)=\frac{0.846}{z}+\frac{1.29}{z^2}+\frac{1.20}{z^3}+\frac{1.01}{z^4}+\ldots \qquad (24)$$

The modified z-transform $G_{Tx}(Z)$ is taken term by term from $G(z)$ by use of the transform pairs of Table II. When this is done with $T_x=0.02$ second, the result is

$$G_{0.02}(z)=1.34\left(\frac{z}{z-1}-\frac{\epsilon^{-0.2}z}{z-\epsilon^{-1}}\right)=\frac{0.242z^2+0.605z}{(z-1)(z-0.368)} \qquad (25)$$

The z-transform for the output 0.02 second after the sampling instants is

$$C_{0.02}(z)=E(z)G_{0.02}(z)=\frac{0.242z^3+0.605z^2}{z^3-1.522z^2+0.890z-0.368} \qquad (26)$$

The first few terms of the solution are placed in evidence by dividing the denominator into the numerator as

$$C_{0.02}(z)=0.242+\frac{0.973}{z}+\frac{1.265}{z^2}+\frac{1.149}{z^3}+\ldots \qquad (27)$$

These steps are repeated with $Tx=0.06$ second. $G_{0.06}(z)$ is

$$G_{0.06}(z)=1.34\left(\frac{z}{z-1}-\frac{\epsilon^{-0.6}z}{z-\epsilon^{-1}}\right)=\frac{0.605z^2+0.243z}{(z-1)(z-0.368)} \qquad (28)$$

$C_{0\cdot 06}(z)$ becomes

$$C_{0.06}(z)=\frac{0.605z^3+0.243z^2}{z^3-1.522z^2+0.890z-0.368} \qquad (29)$$

which when expanded is

$$C_{0.06}(z)=0.605+\frac{1.165}{z}+\frac{1.237}{z^2}+\frac{1.063}{z^3}+\ldots \qquad (30)$$

The $C_{0.02}(z)$ points are shown in Fig. 4. by x-marks and the $C_{0.06}(z)$ points by dots. The curve of Fig. 4 is obtained by using the impulse-response approach.[3]

By the technique outlined in the preceding paragraphs, as many points as desired can be found between sampling periods. The denominator of $C_{Tx}(z)$ remains the same; however, for each value of Tx a new numerator must be found.

Continuous Systems which Coincide with Sampled-Data Systems at the Sampling Instants

The inverse z-transform of a function can be found by an infinite series expansion in descending powers of z or by a partial fraction expansion with the corresponding time functions identified by use of Table I. This latter type of expansion, when combined with methods explained in this section, adds a great deal to the designer's circumspection of the response of sampled-data systems. This method is demonstrated with an example using the circuit of Fig. 3(A) in which $G(s)$ and T are given by equation 20. When a unit step function is applied to the input, the $C(z)$ for this system is given by equation 23 and the denominator of this equation can be factored as

$$C(z)=\frac{0.846z^2}{(z-1)(z-0.261+j0.547)(z-0.261-j0.547)} \qquad (31)$$

$[C(z)]/z$ can be expanded in partial fractions and the result multiplied by z to obtain $C(z)$ as

$$C(z)=\frac{z}{z-1}-\frac{0.51\underline{/-11^\circ}z}{z-0.261+j0.547}-\frac{0.51\underline{/+11^\circ}z}{z-0.261+j0.547} \qquad (32)$$

The time function $c(t)$ corresponding to $C(z)$ can be found by use of Table I as

$$c(t)=1-0.51\underline{/-11^\circ}\epsilon^{-(5+j11.3)t}-0.51\underline{/+11^\circ}\epsilon^{-(5-j11.3)t} \qquad (33)$$

Equation 33 can be written as

$$c(t)=1-\epsilon^{-5t}(\cos 11.3t-0.194\sin 11.3t) \qquad (34)$$

The $c(t)$ of equation 34 can be plotted but it gives the correct output for the sampled-data system at the sampling instants only.

A study of Table I reveals that each z-transform not only has a function of time associated with it but also a Laplace transform. Actually there are an infinite number of Laplace transforms corresponding to each z-transform because the transformation $z=\epsilon^{sT}$ is multivalued in going from the z-plane to the s-plane.

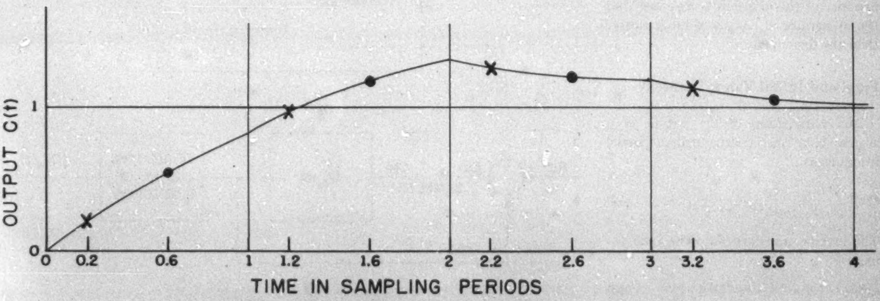


Fig. 4. Response of the system of Fig. 3 showing points for Tx = 0.2T and Tx = 0.6T

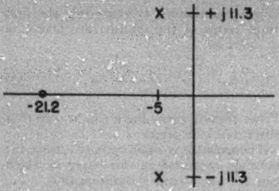


Fig. 5 (left). Pole-zero configuration of equation 37

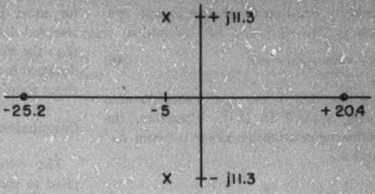


Fig. 6 (right). Pole-zero configuration of equation 38

The convenient Laplace transform to use is the one containing its poles in the primary strip in the s-plane. The Laplace transform associated with equation 32 is

$$C(s)=\frac{1}{s}-\frac{0.51\angle -11^\circ}{s+5+j11.3}-\frac{0.51\angle +11^\circ}{s+5-j11.3} \qquad (35)$$

which is, of course, the Laplace transform of equation 33. In other words, equation 35 is the Laplace transform of a continuous system having an output which takes on the same set of values at the sampling instants as does the output of the corresponding sampled-data system. After terms are combined, equation 35 becomes

$$C(s)=\frac{7.2(s+21.2)}{s(s+5+j11.3)(s+5-j11.3)} \qquad (36)$$

The pole-zero configuration of the continuous system is placed in evidence as

$$\frac{C(s)}{R(s)}=\frac{7.2(s+21.2)}{(s+5+j11.3)(s+5-j11.3)} \qquad (37)$$

and is shown in Fig. 5.

This technique enables the designer of sampled-data systems to use all the experience he has gained and all the material now available in the literature concerning the step-function response of continuous systems having various pole-zero configurations. This technique also eliminates the need for duplicating this work for sampled-data systems.

If a unit ramp function is applied to the input of this system with $t=0$ coinciding with a sampling instant, the same procedure can be followed in finding the pole-zero configuration for the continuous system which coincides with the sampled-data system at the sampling instants. The pole-zero configuration is placed in evidence as

$$\frac{C(s)}{R(s)}=\frac{-0.296(s+25.2)(s-20.4)}{(s+5+j11.3)(s+5-j11.3)} \qquad (38)$$

and is shown in Fig. 6.

These two examples point out several facts. No matter what input is applied to a sampled-data system, there exists a continuous system which coincides with the sampled-data system at the sampling instants. However, there is a different continuous system for each input. Each of these continuous systems has the same poles but different zeros.

Error Coefficients in Terms of z-Transforms

For the single-loop unity feedback sampled-data system of Fig. 7(A), the error coefficients can be defined as

$$K_v=\frac{1}{\text{steady-state error with } r(t)=t} \qquad (39)$$

and

$$K_a=\frac{1}{\text{steady-state error with } r(t)=\frac{1}{2}t^2} \qquad (40)$$

where

K_v = velocity error constant and
K_a = acceleration error constant

When $r(t)=t$, $R(z)=(Tz)/(z-1)^2$ and $E(z)$ is given by

$$E(z)=\frac{R(z)}{1+G(z)}=\frac{Tz}{(z-1)^2[1+G(z)]} \qquad (41)$$

The final value theorem can be used to find the steady-state error as

$$e_{ss}=\lim_{z\to\infty}\frac{T}{(z-1)[1+G(z)]}$$

$$=\frac{T}{\lim z\to1\,[(z-1)G(z)]} \qquad (42)$$

K_v is therefore

$$K_v=\frac{1}{e_{ss}}=\frac{1}{T}\lim_{z\to1}[(z-1)G(z)] \qquad (43)$$

In a similar manner, K_a can be shown to be

$$K_a=\frac{1}{T^2}\lim_{z\to1}[(z-1)^2G(z)] \qquad (44)$$

The error coefficients for the sampled-data system of Fig. 7(A) can be compared with the error coefficients for the continuous system of Fig. 7(B) where the two $G(s)$ functions are the same. It should be noted that $G(z)$ and $G(s)$ are the z-transform and Laplace transform of the same $g(t)$. It should also be noted that

$$\lim_{z\to1}(z-1)G(z)=g(0) \qquad (45)$$

because this is simply the final value theorem being applied to $G(z)$. The

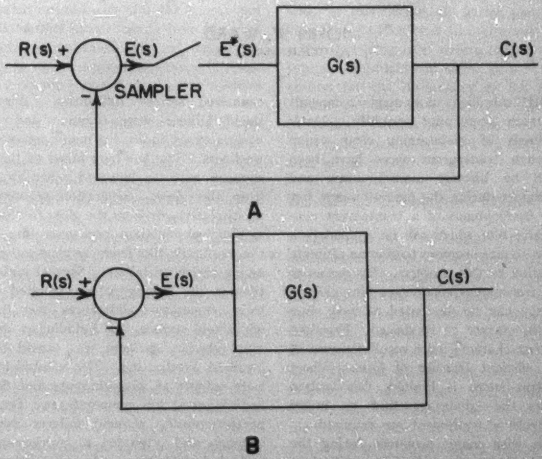


Fig. 7. Sampled-data system and continuous system containing same G(s) function

velocity constant for the continuous system is denoted by K_v^C and is found as

$$K_v^C = \lim_{s \to 0} sG(s) = g(0) \qquad (46)$$

because this is the final value theorem being applied to $G(s)$. Therefore, the following relationship exists between K_v^C and K_v

$$K_v = \frac{1}{T} K_v^C \qquad (47)$$

By a similar development it can be shown that

$$K_a = \frac{1}{T} K_a^C \qquad (48)$$

Without critical examination, equations 47 and 48 seem to indicate that, if a sampler with a sampling period T less than one is added to a continuous system, the error constants are increased resulting in an improved system response. The fallacy in this reasoning results from failure to recognize the fact that K_v^C and K_a^C of equations 47 and 48 respectively are not the K_v^C and K_a^C that would be used in a continuous system to meet the same specifications such as maximum overshoot for which the sampled-data system is designed. It can be shown that for most purposes the response of a sampled-data system is less satisfactory than the response of a continuous system designed to meet the same maximum overshoot requirement.

Conclusions

The z-transformation theory as applied to sampled-data systems is still in its infancy. The techniques presented in this paper add to the designer's circumspection of sampled-data systems.

Appendix. Example Using Notation of the Paper

Given the circuit of Fig. 1 with

$$G(s) = \frac{s}{s+1} = 1 - \frac{1}{(s+1)} \qquad (49)$$

The z-transformation will first be taken improperly to dramatize this situation. The z-transform of equation 49 is

$$G(z) = 1 - \frac{z}{z - \epsilon^{-T}} = -\frac{\epsilon^{-T}}{z - \epsilon^{-T}} \qquad (50)$$

If $r(t)$ is a unit step function, $C(z)$ is

$$C(z) = R(z)G(z) = -\frac{z\epsilon^{-T}}{(z-1)(z-\epsilon^{-T})} \qquad (51)$$

When equation 51 is expanded, the first three terms of the solution are placed in evidence as

$$C(z) = 0 - \frac{\epsilon^{-T}}{z} - \frac{\epsilon^{-T}(1-\epsilon^{-T})}{z^2} - \ldots \qquad (52)$$

From physical reasoning it can be seen that this result is incorrect.

The example is again repeated using the notation suggested in this paper in which case the z-transform of equation 49 is

$$G(z) = 1 - \frac{z}{z - \epsilon^{-T}} V \qquad (53)$$

and the z-transform for the output is

$$C(z) = R(z)G(z) = \frac{z}{z-1} - \frac{z^2}{(z-1)(z-\epsilon^{-T})} V \qquad (54)$$

The first term is not multiplied by V and therefore signifies impulses while the second term is multiplied by V and therefore signifies values.

References

1. THE ANALYSIS OF SAMPLED-DATA SYSTEMS, J. R. Ragazzini, L. A. Zadeh. *AIEE Transactions*, vol. 71, pt. II, Nov. 1952, pp. 225–32.

2. THE ANALYSIS AND DESIGN OF SAMPLED-DATA SYSTEMS, G. V. Lago. *Ph.D. Thesis*, Purdue University, Lafayette, Ind., May 1954, pp. 7–27.

3. THE DESIGN OF SAMPLED-DATA FEEDBACK SYSTEMS, G. V. Lago, J. G. Truxal. *AIEE Transactions*, vol. 73, pt. II, Nov. 1954, pp. 247–53.

Aircraft Windshields Heated by Means of Transparent Conductive Films

JOHN W. WARD
ASSOCIATE MEMBER AIEE

THE safe flight of an airplane depends upon good pilot visibility. Many methods of maintaining clear vision through transparent areas have been used by aircraft manufacturers and operators during the past 10 years, but the development of a transparent conductive film which can be applied to a glass surface promises to offer an ultimate solution to this problem. By means of the transparent conductive film, electric energy can be dissipated as heat close to the surface to be deiced. Excellent deicing characteristics result because of the efficient transfer of heat. Where cockpit space is limited, this system offers the advantage that no bulky controls or equipment are required.

As with many products during the development stages, the first installations of electrically heated windshields encountered serious difficulties. Windshield failures were common and of various types, indicating many points of weakness. The bus bars failed at their connection to the film and pulled away from the glass. Separation of vinyl plastic interlayer from the glass resulted in visual obstructions, and separation of the conductive film from the glass caused arcing and glass failure. Normal variations in the heating pattern resulted in wide temperature differences over the windshield surface, and bubbling of the vinyl plastic interlayer was caused by localized overheating. The windshields were subject to glass-chipping and delamination at low temperatures. Temperature-sensing element failures were common and often led to overheating and windshield failure. In spite of these difficulties, the deicing performance of the windshields was found to be excellent and, because of the inherent advantages offered by this method, research and development were undertaken by the glass manufacturers and the Boeing Airplane Company to attempt to determine the causes of the service troubles and develop suitable remedies.

It is the purpose of this paper to describe this research and development and to discuss the resulting improvements in design, quality control, and operating procedures which have combined to make electrically heated windshields satisfactory for use under aircraft environmental conditions. In addition, the more important considerations in the design and application of aircraft electric heated windshield are discussed.

Description of the Windshield

To be able to discuss the problems encountered with electrically heated

---

Paper 54-525, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted October 24, 1952; made available for printing September 9, 1954.

JOHN W. WARD is with the Boeing Airplane Company, Seattle, Wash.

windshields, it is desirable to describe briefly the construction and characteristics of a typical windshield. A cross-sectional view of a typical aircraft electrically heated windshield is shown in Fig. 1. The inner pane of the windshield is required to carry the pressurization load; it overhangs or bears upon the aircraft structure and is tempered or heat-strengthened to have the highest possible bending modulus of rupture. The inner pane thickness is chosen to provide a factor of safety consistent with the structural requirements of the airplane.

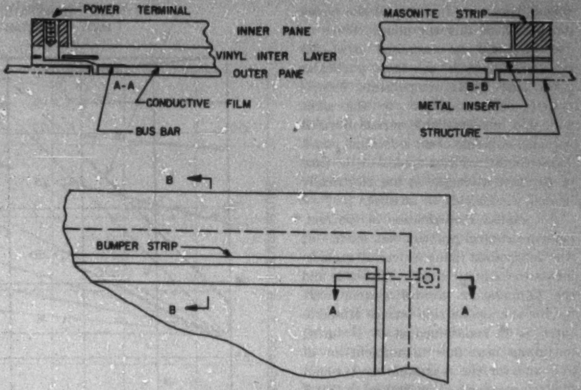


Fig. 1. Typical section of electrically heated windshield

To impart to the windshield a shatterproof characteristic, a vinyl (polyvinyl butyral) interlayer is incorporated. The vinyl interlayer absorbs the energy of shock loads and, in the event of a glass failure, will prevent most of the particles from leaving the windshield. The thickness of the vinyl interlayer may be chosen so that the interlayer is capable of supporting the entire pressure load after both panes have failed. Thus, the vinyl interlayer in the properly designed windshield serves two important functions: it makes the windshield shatterproof and capable of absorbing large shock loads at certain temperatures, and it provides protection against disastrous explosive decompression which otherwise might occur in the event of a glass failure

Depending upon the particular installation, metal inserts may be imbedded in the edges of the vinyl at the point of attachment of the windshield to the aircraft frame. The metal inserts prevent excessive deformation of the vinyl because of frame clamping action and they provide a positive attachment of the vinyl to the aircraft structure.

The transparent conductive film is applied to the inner surface of the outer pane, next to the vinyl interlayer. The film is placed on this surface so that it will be protected from mechanical damage and be insulated electrically, although, from a functional standpoint, more efficient deicing operation would be realized if it were placed on the exposed outer surface. The film is composed of stannic oxide with traces of other materials added to produce desired characteristics. The film is applied when the glass has been heated in an electric furnace to a temperature near its softening point. The glass panel is usually held by tongs while in the furnace, and the tong marks remain after the process has been completed which often result in optical distortion in the adjacent areas. Immediately after the film has been applied, the pane is tempered by chilling it rapidly in a controlled air blast.

The film is approximately 20 millionths of an inch thick and is as hard as the glass itself. It is transparent to directly transmitted light but, by reflected light, a characteristic color pattern determined by the thickness variation is visible. The film is not affected by low temperatures and will withstand temperatures up to about 700 degrees Fahrenheit (F). The resistivity may be varied between the limits of 70 to 300 ohms per square for windshield applications. A nominal resistivity of 120 ohms per square has been standard for aircraft electrically heated windshields in the past, but windshields have recently been designed with the resistance specified within the 70- to 300-ohm-per-square range to result in a desired power dissipation at voltages of 120 to 208 volts.

Electric connections to the film are made by means of bus bars located along opposite or parallel edges. The bus bars consist of a fired-on metallic coating over which the transparent conductive film is applied. An additional painted or sprayed-on metallic layer is applied over the conductive film and to this layer a copper braid is soldered. The copper braid is imbedded in the vinyl and is attached to a terminal. The bus bar is less than 0.005 inch thick and from 1/4 to 3/8 inch wide. The electric conductivity of the construction is good and no appreciable voltage drop takes place along its length which might reduce the amount of current received by the areas furthest removed from the connection to the bus.

The outer pane serves as a cover plate and is not relied upon to contribute to the rigidity of the windshield. It does not overhang the structure as does the inner pane, yet, when the entire windshield bends with pressure loading, the outer pane must be capable of assuming this deflection without failure. The flow of heat from the conductive film to the outer surface is through the outer pane. The thermal gradients caused by this flow of heat produce in the glass bending stresses which are additive to the pressurization stresses. The bending strength of the outer pane must therefore be great enough to enable it to withstand these combined stresses and, for this reason, the outer pane is also tempered or heat-strengthened.

The foregoing brief description of the typical windshield will serve as a background for the discussions of the important problem of temperature control to follow. Additional details of the windshield construction will be covered later in the paper.

Windshield Temperature Control

The failures of windshields on the first installations indicated that the temperature controls had allowed film temperatures to exceed safe limits. It was then discovered that the power dissipation on the conductive film varied widely above and below the average and, furthermore, the heat pattern was found to be different for each panel. One temperature-sensing element was used to control the temperatures of several windshields. The location of the element in one of the windshields was fixed so that the power supplied to all of the controlled windshields became a function of the power dissipation at this one point. Since no control could be exercised over the power dissipated at that particular point, the whole temperature control system became subject to the same wide and uncontrollable variation which was characteristic of an individual panel. These conditions were typical at the time of the first attempts to use electrically heated windshields in aircraft.

A complete re-evaluation of the temperature control problem was made. It was determined that the over-all requirements to be met by the windshields and the temperature control system were for the average outside surface temperatures to be maintained at 40 F during maximum heat flow to ambient air at 0 F and for the maximum or hot-spot film temperatures to be limited to 160 F during operation at minimum ambient temperature. In other words, the windshields must be deiced without overheating.

The following sections describe the heat-flow studies which resulted in a simple means of relating power dissipation and temperature at a point; the development of a test method by which power dissipation could be measured; and the development of temperature control system components to provide reliable and accurate temperature control.

Power Constants

The flow of electric current and the flow of heat are analogous in a number of respects. Steady-state heat flow for simple heat circuits follows a sort of Ohm's law in which temperature differences are equivalent to voltages and thermal and electric resistivities limit the flow of heat and current in similar ways. The thermal counterpart of Ohm's law becomes

$$Q=\frac{\Delta T}{R}$$

where

Q = rate of heat flow, British thermal units per foot^2 per hour (Btu/ft^2/hr)

ΔT = temperature difference in F between two points on a line parallel to the heat flow

R = thermal resistance in [(ft^2)(hr)(F)]/(Btu) between two points parallel to the heat flow

The window heat-flow problem can be explained most easily by first making the following simplifying assumptions.

1. Assume that no appreciable heat flows to the cabin side of the windshield. This is true under normal flight conditions.

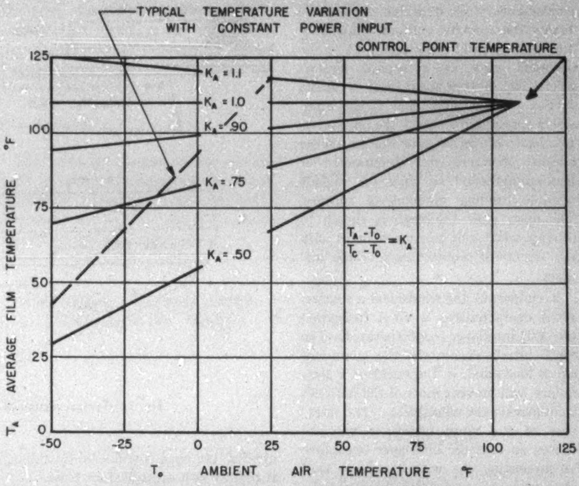


Fig. 2. Average film temperature versus ambient air temperature and power constant K_A for constant control-point temperature and variable power input

2. Assume that h_0, ($h_0=1/R_F$ where R_F is the thermal resistance of the air film) the heat-transfer coefficient, over the outer glass surface is uniform.
3. Assume steady-state heat flow.

It can be shown that at each of two points on the film, which can be called H and C to designate the hot spot and control point respectively, the total temperature rise above ambient is equal to the temperature drop at each point caused by Q_H and Q_C flowing through R_G plus R_F or

$$\Delta T_H = Q_B(R_G+R_F)$$

$$\Delta T_C = Q_C(R_G+R_F)$$

where

R_G and R_F = glass and air film resistances respectively

Q_H and Q_C expressed in terms of Btu/ft^2/hr are related to the watts dissipated per ft^2 by a factor of 3.41 or (3.41) watts/ft^2 = Btu/[(ft^2)(hr)]

Power dissipation in watts/ft^2 at points C and H will be designated P_C and P_H

The resistivity at a point on a particular panel does not change with temperature or stress and, therefore, the relative current distribution in the film will not be changed during normal operation. At any point, the current will always be proportional to the bus-to-bus voltage. The ratio of the power-per-unit area dissipated at any two points will, therefore, be a constant and it can be stated that $P_H/P_C=K$.

Substituting this relationship into the equations for the total temperature drops, it is seen that

$$K=\frac{P_H}{P_C}=\frac{Q_H(R_G+R_F)}{Q_C(R_G+R_F)}=\frac{\Delta T_H}{\Delta T_C}=\frac{T_H-T_0}{T_C-T_0}$$

where T_0 is the ambient air temperature.

This series of very simple equations show that the ratio of the temperature rises above ambient at two points is a constant and that this relationship is independent of the ambient temperature and the heat-transfer coefficient. Because no satisfactory method has been developed for calculation of the heat-transfer coefficient h_0, it is indeed fortunate that this variable does not enter into the determination of the relative temperature rises. Expressions similar to those in the foregoing can also be developed for more general steady-state heat-flow cases where heat loss to the cabin is considered and groups of windshields are controlled by one temperature-sensing element.

Power constants defined as follows have been found to describe the important windshield heat-pattern characteristics

$$K_A=\frac{\text{average power}}{\text{power at the control point}}=\frac{P_A}{P_C}$$

$$K_H=\frac{\text{power at the hot spot}}{\text{power at the control point}}=\frac{P_H}{P_C}$$

$$K_M=\frac{\text{average power}}{\text{power at the hot spot}}=\frac{P_A}{P_H}$$

The power constant K_A indicates the

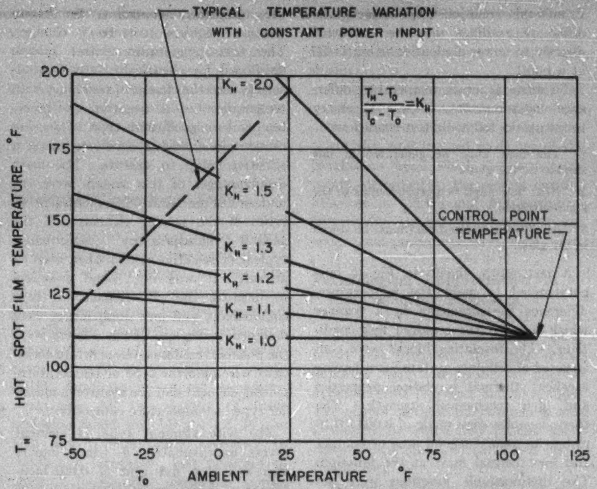


Fig. 3. **Hot-spot film temperature versus ambient air temperature and power constant K_H for constant control-point temperature and variable power input**

relative rises above ambient of the average and control-point film temperatures. Since it is the average temperature which is of concern in deicing calculations, the constant K_A can be said to describe the deicing characteristics of the windshield. The control-point temperature is normally maintained constant by the temperature control system, and the temperature rise at the control point is the difference between the ambient and control-point temperatures. Knowing the power constant K_A for a given windshield, the average temperature of the conductive film for any ambient temperature can be readily determined. Curves of calculated average film temperatures versus ambient air temperature are shown in Fig. 2 for various values of K_A.

The power constant K_H indicates the relative rises above ambient of the hot-spot and control-point film temperatures. The relationship between hot-spot film temperatures and ambient air temperature is shown in Fig. 3 for various values of the constant K_H. A most difficult concept to grasp is presented in this figure, where it is predicted that the temperatures at points having a power dissipation greater than that at the control point will rise as the ambient air temperature decreases. Normally it is expected that a low ambient air temperature should be accompanied by lowered glass temperatures; however this is not the case. From the curves of Fig. 4 it can be seen that the constant K_H for a windshield to be operated at −65 F with a control temperature of 110 F must be 1.3 or less if the hot-spot temperatures are to be limited to 160 F.

The constant K_M represents the ratio of average to maximum power dissipation and, therefore, is a direct indication of over-all heat-pattern quality. The constant K_M is always equal to K_A/K_H and is the only one of the three constants which does not change with the sensing-element location.

It has been shown that by means of the power constants the important properties of a given windshield heating pattern are described. It was logical then to use the power constants as the basis for production quality control of the conductively coated panels. This was done and limiting values for each power constant were specified. To realize the greatest advantage from the use of the power-constant quality control method, it was necessary for certain areas of the panel to be designated within which the sensing element could be located. It can be seen by studying the nature of the constants K_A and K_H that their values are dependent upon the power dissipation at the particular point chosen as the control point.

By moving the sensing element to the hot spot, the constant K_H becomes unity and the constant K_A becomes smaller, maximum temperatures are limited to the control-point temperature, and the average temperature is lowered resulting in poorer deicing performance. As the sensing element is moved into an area of average power dissipation, the constant K_A approaches unity and the constant K_H increases, the average film temperatures will be maintained constant resulting in good deicing performance, while hot-spot temperatures will be allowed to rise above the control-point temperature and overheating may result. A compromise between these two extremes may be achieved by proper choice of power-constant limits and by allowing

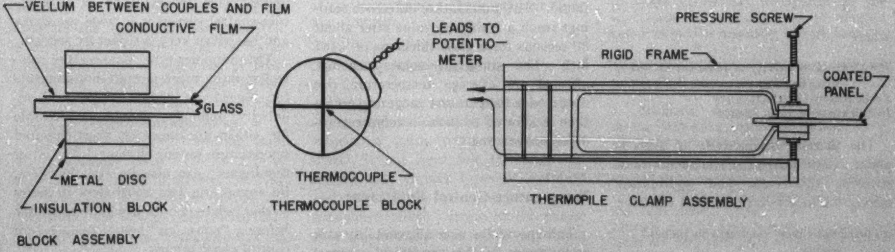


Fig. 4. **Test equipment for measuring temperature differences**

sufficient latitude in the location of the temperature-sensing element.

Typical power-constant limits for an aircraft windshield are as follows:

$K_A = 0.75$ or greater

$K_H = 1.3$ or less

$K_M = 0.65$ or greater

A windshield having perfectly uniform power dissipation would have all three power constants equal to unity. The power-constant system of heating-pattern quality control is sufficiently flexible so that manufacturing improvements can be reflected in power-constant limits which become closer to unity as the heating pattern becomes more uniform.

Power-Constant Measurements

In the actual sequence of events, the power-constant concepts were developed and power-constant limits which would result in the desired windshield characteristics were fixed but no entirely satisfactory method of measuring power dissipation at a point was known. As a first attempt, stabilized temperature-rise measurements were made in still air and the stabilized temperature rise was assumed to be proportional to the power dissipation. It was soon discovered that this method was subject to many inaccuracies. During subsequent investigations, an experiment was conducted on an unlaminated conductively coated pane which indicated that thermocouples, placed as shown in Fig. 4, measured a temperature difference which was proportional to power input, independent of the temperature rise of the glass, and independent of time after a brief period of stabilization. It was then found that the temperature difference depended only upon three factors: the power dissipated, the thermal conductivity of the glass, and the glass thickness. This relationship is expressed as follows

$$\Delta T = \frac{RIl}{2}$$

where

R = thermal resistivity of glass (hr)(ft²)(F)/(Btu)(inch)

I = heat flow (Btu)/(ft²)(hr)

l = thickness of glass, inches

The thermal conductivity of glass is about 7[Btu(inch)]/[(ft²)(hr)(F)]. Substituting this value and appropriate conversion factors into the first equation

$$\Delta T = 0.2433 \text{ (glass thickness in inches) } \frac{\text{(watts)}}{\text{ft}^2}$$

From this relationship, temperature-difference readings may be converted directly to power dissipation in watts/ft² at a point.

To make accurate temperature-difference measurements, it was necessary to set up the following test conditions:

1. The heat must be generated at one surface of the glass.
2. Heat flow must be perpendicular to the glass surfaces.
3. All of the generated heat must be stored in the glass.

A test setup, shown in Fig. 4, was constructed to make the temperature-difference measurements in a manner which satisfied the required test conditions. The insulating blocks were constructed of cellulose acetate foam having excellent thermal insulation properties and high mechanical strength. The thermocouples were made of 0.003×0.10 ribbon to assure a good area of contact and low thermal mass at the junction. The thermocouple junctions were accurately centered on the insulation blocks and a clamp assembly was constructed to hold the insulation blocks in proper alignment. To protect the film from abrasion and to insulate the thermocouple electrically, a layer of vellum was inserted between the thermocouple block and the film. This added thermal resistance did not introduce an error as long as an equal thickness of vellum was also inserted between the thermocouple and the glass on the uncoated side. A millivoltmeter to measure the thermopile output and a wattmeter to measure power input were also required for making power-constant measurements.

To measure the power constants of a conductively coated panel, first the average power input from a wattmeter reading corrected for the heated area to give watts/ft² is determined. The power at the hot spot and control point are then determined by measurement of temperature differences. From these three power measurements, the power constants are found. The temperature-difference readings reach a stabilized value after about 30 seconds for a glass thickness of 0.187 inch. The stabilized values are not affected by average temperature rise over a wide temperature range and ample time is allowed to make a temperature-difference reading.

Temperature-Control Equipment

To improve the over-all reliability and performance of the windshield system, it was found that certain deficiencies of the temperature-control system components required corrective changes. The first temperature-control system developed for electrically heated windshields used the change of resistance with temperature of a negative resistance-temperature coefficient type of sensing element in a bridge circuit to cause a sensitive relay to operate. The main disadvantages of this system were the undesirable mechanical and thermal properties of the sensing element and the general unsatisfactory performance typical of sensitive relays when used in aircraft. It was not found practical to improve the performance of these components, and new components were designed to replace them. These were the positive resistance-temperature coefficient wire filament type of temperature-sensing element and the electronic amplifier type of temperature controller.

The wire filament type of element covers approximately a 1-inch square area and is a flat grid of 0.001-inch-diameter high resistance-temperature coefficient wire. The element is nearly invisible when it is installed in the windshield and its small thermal mass does not introduce a thermal lag error. The filament can be located within 0.020 inch of the conductive coating and, because it is thin, does not restrict the minimum vinyl thickness as did the thicker negative resistance-temperature coefficient type of sensing elements. The resistance calibration of the element can be controlled closely. The self-heat error can be kept to a minimum by limiting the voltage impressed across the sensing element. The wire filament type of sensing elements are easily installed and are not subject to failure caused by installation stresses.

The temperature-sensing element is connected as one leg of a bridge circuit. The output of the bridge is fed into an amplifier which operates a control relay to turn on or off the windshield power as required to keep the bridge circuit balanced. This general type of temperature-control system is in wide use today and has given very satisfactory service.

The amplifier-type temperature controller must meet aircraft requirements for ruggedness, light weight, and reliability and also should be fail-safe for either an open- or short-circuited temperature-sensing element. The latter requirement can also be met, in effect, by connecting two controllers in series so that failure of both would be necessary before a "power on" loss of temperature control could occur.

The development of the power-constant

method of quality control, the wire filament type of temperature-sensing element, and the electronic temperature controller have satisfactorily fulfilled the temperature-control system requirements; however, these developments alone did not furnish the complete solution to the windshield failure problem.

Laboratory Operational Tests

In addition to the temperature-control problems encountered, there existed real problems of a structural nature. The flow of heat and the resulting thermal gradients and thermal stresses were found to influence adversely the structural characteristics of the windshields. Several conditions resulting in windshield deterioration caused structural weakness. All the aspects of these problems could not be studied in the field and it was found necessary to conduct a laboratory test program to attempt to simulate conditions during which failures were known to occur. Out of this laboratory test program developed standard simulated operational tests which are used as a basis for the comparison of windshield qualities and as a means of evaluating construction improvements. The most severe flight condition is simulated by these tests, i.e., maximum pressure deflection, maximum power input, and maximum normal sensing element temperature, at minimum ambient temperature. During this condition the hot spots are hottest, the cold spots are coldest, the pressure bending stresses are maximum, and the stresses produced by heat flow are maximum. Successive tests are run at higher pressures to determine the approximate factor of safety of the windshield when it is both heated and pressurized.

The simulated operational tests are conducted in a cold chamber in which the ambient temperature can be closely controlled. The test windshield is mounted in a frame constructed to duplicate the mechanical characteristics of the aircraft structure and the frame is mounted on a box which can be pressurized. A high-velocity fan directs an air stream over the outside windshield surface to conduct the heat away.

The types of failures resulting from the operational tests were similar to those of windshields in service. As long as this parallel in experience existed, the laboratory test results could reasonably be expected to be representative of service conditions. Corrective measures were worked out to improve performance during laboratory tests and, if the measures were successful, they were also applied to the production windshields.

Analysis of laboratory operational test results has supported the validity of the power-constant method of temperature calculation and has led to many of the conclusions upon which design requirements, described in the section entitled "Factors in Design and Application of Electrically Heated Windshields," and operational procedures have been based.

Technological Improvements in Manufacturing Processes

Concurrent with the laboratory operational test program, the glass manufacturers were attempting to improve manufacturing processes and techniques. The effects of these improvements were evaluated by subjecting typical specimens to the standard simulated operational tests.

It is not within the scope of this paper to describe in detail these improvements; nevertheless, they have contributed greatly to the development of the electrically heated windshield. Some of the more important advances which resulted in improved performance during the operational tests were as follows:

1. Bus bar design improvements have eliminated failures at these locations.
2. The glass temper, bending modulus of rupture, obtainable in a conductively coated panel has been increased through the improvement of manufacturing methods.
3. Control of the uniformity of film resistivity has been improved.
4. Quality control methods have been developed to detect glass and film irregularities.
5. Windshield optical properties have been improved by careful selection of tong locations and control of warpage during the tempering operation.
6. The adhesion of the film to the glass and of the vinyl to both the film and the glass have been improved by careful control of the many factors which influence these characteristics.

Operating Procedures

The operating procedures for electrically heated windshields have been the subject of many controversial discussions. It has been felt by many observers that certain types of warm-up and cool-off cycles are required to prevent windshield failure. It was for this reason that an attempt was made to obtain from an analysis of test results the requirements for an optimum operating cycle to prevent thermal shock windshield failures. The conclusion reached after a thorough study of the data was that the most severe operating condition was that which required continuous maximum power output from the windshield while pressurized at minimum ambient air temperature. It was concluded that windshields which failed from thermal shock during rapid warm up would have ultimately failed if they had been required to deliver maximum power at minimum ambient temperatures. The first windshields survived as long as they did primarily because during a large percentage of the time in operation the ambient temperatures were moderate and the input power required was low. In this case, an automatic warm-up cycle would have extended the service life of the poor-quality windshield but would not have prevented failure during the more critical steady-state operating condition. On the other hand, it was concluded that automatic warm-up controls could do little to extend the service life of a windshield already capable of withstanding the critical steady-state condition. For these reasons, it was felt that, to reach an ultimate solution to the windshield failure problem, emphasis should be placed upon development of the windshield rather than upon the development of automatic temperature controls.

The effect of the correct operating procedure on windshield service life is not to be underestimated however, and the study to determine an optimum operating cycle resulted in the following recommended procedure:

1. An available power input to the windshield of between 1,000 and 1,500 Btu/ft^2/hr is recommended for all operating conditions except those which require greater heat for maintaining clear vision.
2. An available power input to the windshield between 2,000 and 3,000 Btu/ft^2/hr is recommended for use during extreme icing conditions.
3. It is recommended that low heat (1,000 to 1,500 Btu/ft^2/hr) be applied to the windshield for approximately 15 minutes before switching manually to high heat.

The following advantages will be realized by using this procedure:

1. Adequate deicing will be afforded during light icing conditions without using the maximum value of input power.
2. Adequate defogging of inside windshield surfaces during the high-altitude cruise condition will be accomplished by the low heat input. Icing conditions at high altitude and at low ambient air temperatures are very rare and, under these conditions, defogging will be satisfactory even though outside windshield surface tem-

peratures fall below 40°F. The most critical operating condition was found to be the steady-state pressurized, low ambient temperature, high power input condition. Since the greatest part of a flight is spent at high-altitude cruise, it can be seen that by reducing the power input during this condition extended periods of subjection to maximum stress can be avoided. The intended function of the windshield will still be accomplished.

3. The need for complicated automatic controls are eliminated and little or no additional burden is placed upon the flight personnel. Insufficient power input is easily distinguished and necessitates the operation of one switch to make maximum power available to the windshields.

4. Even with low heat input during most take-off and climb conditions, the vinyl temperatures will be normal and the impact resistant properties of the windshield will not be diminished.

5. The on-off power demand may impose severe service on electric equipment, especially when several large windshields are connected together. Low power input will result in a slower cycling rate and a lower power required during the on period so that the severity of the duty on electric components will be lessened.

To the author's knowledge, no windshield failures have been reported during a cool-off period; however, it is possible that a properly controlled cool-off period could prevent some types of windshield deterioration. When the power input is reduced to zero, the temperatures of the windshield will approach ambient exponentially. It is possible to "lock in" vinyl stresses which will not be equalized by vinyl flow at temperatures below 50 F even though temperature gradients eventually become zero. To be effective, a cool-off period would have to allow complete equalization of vinyl temperature and stresses above 70 F to that there will be no locked-in stresses when the temperatures drop below 50 F. It is obviously impractical to devise a cool-off cycle meeting this requirement; therefore, the only practical solution to this problem appears to be to make the windshields capable of withstanding the shock or stress which accompanies a normal cool-off.

As an illustration of the improved ability of the windshields to withstand thermal shock, a test specimen typical of the most recent type withstood repeated warm-up and cool-off in an ambient temperature of −65 F with a power input of approximately 3,000 Btu/ft²/hr while sustaining an internal pressure of three times normal or about 20 pounds per square inch. The first windshields of this type were not consistently capable of withstanding warm-up at 2,000 Btu/ft²/hr from 0 F while sustaining no pressure. Results of this nature support the contention that automatic warm-up controls which would have been helpful in extending the service life of a windshield of the first type add little to the service life of a windshield of the latest type.

The foregoing material has described briefly the separate problems which have required major study and development. During these various development stages, minor improvements, limitations, and design requirements have developed which are all of importance because of their contribution to the overall satisfactory performance of the electrically deiced windshield. These factors will be considered in the following.

Factors in Design and Application of Electrically Heated Windshields

Some of the more important design considerations which have resulted from laboratory, manufacturing, and service experience are described in the following.

THERMAL AND ELECTRIC POWER INPUT RELATIONSHIPS

Assuming a nominal conductive film resistivity of 120 ohms per square, the approximate bus-to-bus resistance of a conductively coated panel in ohms will be equal to 120 times the average width (distance between bus bars) divided by the average length (parallel to the bus bars) of the heated area. The windshield manufacturer is normally allowed a tolerance of ±15 per cent of the estimated bus-to-bus resistance for production windshields. It has been the practice to divide this resistance variation into three equal ranges, each designated by a letter or number sandblasted in an appropriate location on the inner surface of the windshield. By providing proper voltage taps on the power transformer, it is possible to supply approximately the same average power input to windshields in each of the three resistance ranges.

The average rate of heat dissipation for a particular windshield expressed in Btu/ft²/hr is determined from the following relationship

$$\text{Btu/ft}^2\text{/hr} = \frac{(3.41)(\text{windshield voltage})^2}{(\text{heated area in ft}^2)(\text{bus-to-bus resistance in ohms})}$$

The maximum power which can safely and usefully be dissipated by a given windshield can be estimated from the following expression

$$\text{Max. average power in Btu/ft}^2\text{/hr} = \frac{[(\text{min kilometers})(160) - 40]\ (7)}{(\text{nominal outer pane thickness in inches})}$$

A power input greater than that given by the foregoing equation can be dissipated only by allowing the average outside surface temperature to drop below 40 F or by increasing the allowable hot-spot film temperature. If the average outside surface temperature drops below 40 F, deicing will not be satisfactory and, if the allowable hot-spot film temperature is increased, there is danger of overheating and attendant vinyl-bubbling or glass failure. Power requirements for deicing based upon thermodynamic calculations for a particular airplane must be co-ordinated with the ability of the windshield to deliver the necessary heat in order to arrive at the most practical system. This precaution is included because it has become evident that calculated heat requirements for deicing generally exceed the value of heat which it is practical to dissipate through the outer pane.

PARTING MEDIUM

A large number of windshield failures in service have been caused by glass-edge chipping which slowly reduces the strength of an outer pane until complete failure occurs. To prevent edge chipping, a construction, termed a parting medium, is used in the locations shown in Fig. 5. The parting medium breaks or weakens the bond between the glass and the vinyl in the edge zones so that stresses cannot be transmitted to the glass. These areas are especially susceptible to chipping because the edges of the glass are cut and ground and furnish many small points for stress concentrations. On the polished flat surfaces, these imperfections are not as pronounced and a greater stress is required to cause the glass to chip.

The ideal parting medium is one which will provide a connection between the vinyl and the glass which is flexible at all temperatures. A parting medium having this characteristic will theoretically prevent glass-edge chipping, moisture penetration, and reduce delamination tendencies. The several parting mediums in general use do not meet all of the requirements for the ideal but they have been successful in preventing glass-edge chipping, which is the most important function. Additional service experience is required before the nature of moisture penetration and delamination problems can be determined.

OUTER PANE THICKNESS

The choice of an outer pane thickness to give optimum performance is made by reaching a compromise between thermal

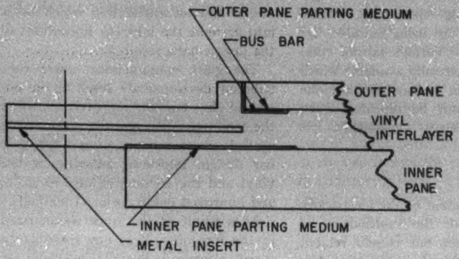


Fig. 5 (above). Typical section showing parting medium location

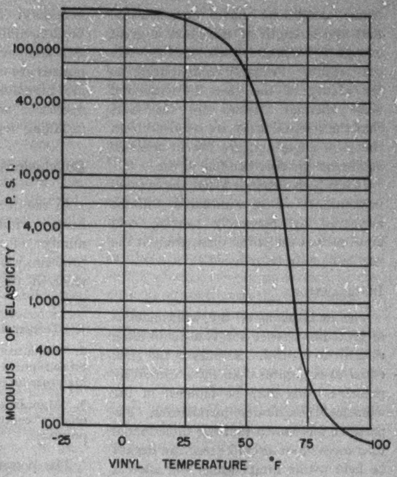


Fig. 6 (right). Modulus of elasticity versus temperature for vinyl plastic

and structural requirements. For a given required heat transfer, the temperature drop through the outer pane is directly proportional to its thickness. Since it is desirable to maintain a given outside surface temperature with as low a vinyl temperature as possible, the temperature drop through the glass should be kept to a minimum by using a minimum glass thickness. The minimum glass thickness which can be used is limited, on the other hand, by structural and optical requirements. Glass less than 0.11 inch thick is difficult to temper without warpage or waviness which influences both optical and structural characteristics. In addition, the maximum temper which can be imparted to a glass panel is somewhat proportional to its thickness. Experience has shown that the outer pane glass thickness which is a practical compromise between the thermal and structural requirements is 0.187 or 3/16 inch. The temper, as indicated by the bending modulus of rupture, which can be imparted to a coated panel of this thickness is between 18,000 and 21,000 pounds per square inch. The temperature drop through the panel with a heat flow of 2,500 Btu/ft^2/hr is approximately 73 F and, to meet the requirements for satisfactory deicing, an average vinyl temperature of (40+73) or 113 F must be maintained.

Vinyl Interlayer

The modulus of elasticity of vinyl varies with temperature, as shown in Fig. 6. The thermal conductivity of vinyl is about 1/5 that of glass. The temperature coefficient of expansion of vinyl is approximately 30 times as great as that of glass. These properties of vinyl are directly responsible for the majority of the service troubles experienced with electrically heated windshields. The high strength at low temperature together with the large temperature coefficient of expansion limits the minimum temperature which the windshield will withstand. The low strength at high temperatures causes the vinyl to deform in accordance with heat-pattern variations. These vinyl deformations produced at high temperature are capable of exerting high stresses, if the windshield is subsequently subjected to low temperatures, which may result in glass failure or delamination. Also, temperature gradients in the vinyl interlayer can cause the vinyl to bend the entire windshield, usually in the same direction as the pressure bending, thus adding to thermal and pressure glass stresses.

The adverse effects of the vinyl interlayer have been found to be somewhat proportional to its thickness. However, the vinyl thickness cannot always be adjusted to result in the minimum tendencies toward deterioration because of the structural requirements which must also be met. A vinyl thickness of 0.060 to 0.100 inch would reduce tendencies toward deterioration and still provide necessary clearance for metal inserts and the temperature-sensing element, but a vinyl interlayer of this thickness would provide very little impact resistance and no blowout protection. Relatively large windshields (approximately 20×20 inches) designed for 6.5-pound-per-square-inch pressure differentials have a 0.25 inch thick vinyl interlayer and larger windshields have been built with an interlayer thickness of 0.40 inch. Tests of these windshields have demonstrated the ability of the vinyl to prevent blowout of the windshield in event of glass failure. Because the value of the interlayer as a structural member is great, structural considerations will generally be the determining factor in the choice of a vinyl thickness.

Windshield Shape

An ideal electrically conductive film is one in which the heat dissipation is uniform over the entire area. With a rectangular heated area, this ideal is achieved by a film having uniform resistivity. Any departure from a rectangular shape requires the film resistivity to be varied in a definite manner in each area to achieve uniform power dissipation. Because of the nature of the coating process, it is very difficult to control the resistivity in given areas. Although much progress has been made in this direction, intelligent choice of windshield shape in the initial design stages will reduce the number of problems involved in achieving a satisfactory heat distribution.

An important means of correcting the shape of the heated area in the case of an irregularly shaped panel is the deletion of conductive film. Laboratory and service experience have shown that deletion of film to eliminate hot spots and improve power constants is a satisfactory way of improving windshield service life and performance. Since the deletion of film in an area will leave that area unheated, it can be expected that thermal stresses in the glass will develop at the boundary between the heated and unheated areas during normal operation. Extensive tests under extreme conditions have demonstrated that the strength of the glass is great enough to withstand these stresses without failure. Probably contributing to the ability of the glass to withstand these thermal stresses are the facts that the stressed areas are removed from the glass edges and are not in areas of maximum pressure-bending stress.

It has been concluded that the over-all performance of a windshield will be improved by purposely causing cold areas so that the power dissipation at the hot spots may be reduced.

INSTALLATION

The installation of the windshield is of great importance and is a detail often difficult to control. Stresses in the glass equal to or greater than those caused by pressure loads may be induced in the windshield by faulty installation. Experience has shown that the dimensional tolerances of an aircraft structure cannot be held within limits which will allow a windshield to be installed without additional adjustment. It has been found necessary to perform a fairing operation which compensates for manufacturing tolerances and provides a support onto which the windshield may be clamped or bolted without inducing stresses. A tolerance of ±0.10 inch on the contour or flatness of the frame is considered necessary. Fairing the frame should be accomplished by using suitable jigs and compounds. Various fairing compounds are commercially available which are permanent in nature so that the operation need not be repeated when installing replacement windshields.

Conclusions

In the development of the electrically heated windshield, the solutions to a number of distinct but closely related problems were required. These problems resolved themselves into the following groups:

1. Temperature control of problems.
2. Application problems involving the establishment of design criteria and operating procedures.
3. Manufacturing problems involving quality control methods, materials, and processes.

The power-constant system of heat-pattern quality control, the method of measuring power constants, and temperature control system components were developed to solve the temperature control problem.

Simulated operational test results were analyzed and operational procedures, installation requirements, and design criteria were developed. An optimum design and an application which takes into account the inherent limitations of the design have resulted.

The glass manufacturers have contributed developments resulting in improved glass temper, improved control of the conductive film resistivity, improved optical characteristics, improved bus bar design, improved adhesion of the vinyl and the conductive film to glass, and improved quality control methods.

As a result of the work summarized in the foregoing, it is now possible to specify electrically heated windshields for a given application with reasonable assurance that the performance will be satisfactory. During the course of this work, specification requirements have continually reflected improvements in design and test methods. The specification has become a standard by which past performance and future improvements may be evaluated; it represents the usable evidence of accumulated test and service experience. It is the desire of many who have contributed to this development that windshield specification requirements and test methods throughout the aircraft industry be standardized. This will make it possible for all aircraft of the future to be made safer by the use of windshields heated by means of transparent conductive films.

The Toronto Subway

J. G. INGLIS
MEMBER AIEE

IN TORONTO, Canada, the last street car and trailer line on the Continent has been replaced by a modern subway system thoroughly integrated with surface transfer lines. Since cars, shops, and track were entirely new, all of the modern features and methods of other systems were studied and adopted where practical. The cars, as well as signals and part of the power supply, telephone system, and cables were purchased in the United Kindgom. Operation of the first few months has already justified the wisdom and courage of those who conceived, planned, and built the system.

General

Toronto became the first city in Canada and the fifth in North America to operate a subway system when, on March 30th, 1954, the Toronto Transit Commission officially opened a $4^{1}/_{2}$-mile underground line, built in 4 years at a total cost of almost $58,000,000. The line connects the downtown district with the north end of the city. It runs under or near the main street and for about two-thirds of its length is underground, the remainder being in open cut. It is the backbone of the system and makes connection with 25 streetcar, trolley-coach, and self-propelled bus lines.

The subway was built to provide a service which already has made about 100,000 persons immune to the traffic paralysis which had been creeping over the downtown district. It has, as was expected, increased the confidence of industrial and commercial interests in the area it serves, and it has raised assessed values in surrounding districts, and demonstrated in yet another city that transit progress is civic progress.

The subway cars were purchased in England from the Gloucester Railway Carriage and Wagon Company in competitive bidding from the United States, Canada, Great Britain, France, and Belgium. Their purchase in Britain also helped to pay for the export of Canadian products such as wheat and cheese which are so necessary to the economy of the nation.

Of the 104 cars purchased, 100 are of steel, whereas the last four utilize aluminum to the maximum practical extent, in order to determine, for the future, if the economics resulting from lighter weight will have justified the additional first cost.

Paper 54-343, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing July 16, 1954.

J. G. INGLIS is with the Toronto Transit Commission, Toronto, Ont., Canada.

Significant dimensions of the cars are as follows:

Car length over bumpers, 57 feet $1^{1}/_{2}$ inches
Car length over body, 55 feet $7^{1}/_{2}$ inches
Car width, 10 feet 4 inches
Car height, 12 feet 0 inches
Truck centres, 38 feet 0 inches
Truck wheelbase, 7 feet 0 inches
Wheel diameter, 30 inches
Door openings per car side, 3
Door opening width, 3 feet 9 inches
Door height, 6 feet 4 inches
Seating capacity, 62

Train Size and Performance

The cars are connected semipermanently in pairs to form 2-car units in order to reduce both initial and maintenance costs. Train size varies with traffic, service being provided at present in the rush hour with 6-car trains operating on a 2 feet 15 inch headway and carrying about 27,000 passengers per hour. A maximum capacity of 40,000 passengers per hour may be carried in either direction by operating 8-car trains on a 2-minute headway.

The 12 stations on the line average 0.4 miles apart and the 1-way running time varies between 14 and 16 minutes depending upon direction of travel and load. This corresponds to minimum and maximum schedule speeds, exclusive of layover of 16.85 and 19.25 miles per hour.

The Car

BODY CONSTRUCTION

The car body is of conventional riveted and welded design with a cab at the coupling end of each unit. The exterior sides and roof are of unusually smooth contour, to permit trains to be operated through the motor-driven car washer. To resist damage resulting from collisions, each car is fitted with an "anticollision pillar," which is a vertical structural member, extending about 18 inches above and below the floor level of each corner. These perform the function of the familiar anticlimber.

Insulation of the car body against temperature and sound has received unusual and successful attention. A very heavy coat of mastic paint covers the complete interior surface of the exterior shell and, in addition, a layer of fabric covers the shell from floor level to window-sill level. Furthermore, the floor covering consists of 1/4 inch of rubber, laid on $1^{1}/_{2}$ inch of cork fastened to the steel floor. These features, together with sound-proofing of the middle wall of the tunnel and of station platform edges, have produced a car which gives an exceptionally quiet ride.

PASSENGER APPEAL

A great deal of attention was given to the appearance and passenger appeal of the car as well as to its quietness of operation. The interior colors are bright and in pleasing contrast, with 2-tone Formica on the sides and ceiling, red plastic-covered form-rubber sets and a gray floor. Interior surfaces which require paint were kept further to a minimum by the use of matching anodized light fixtures and sash, stainless-steel stanchions and Formica-and-glass door barriers. The 24 ceiling-mounted standee grips were made as inconspicuous as possible by turning them to present minimum area to longitudinal view. Also, the number required was reduced by fitting each one for the use of two persons. The advertising card frame extends in a uniform unbroken line at the same height around the complete car interior. The exterior of the car is a bright red with gold stripes, and the roof is black.

DOORS

The doors have several unique features. While conventional, in that there are three double-width door openings per car side, each opening has an aluminum-alloy 2-section sliding door which is not hung from the top but runs on rollers on a steel floor plate. This arrangement gives freedom from the troubles associated with floor grooves. In addition to the usual abutting rubber edges, the trailing section of each door may be forced open about $4^{1}/_{2}$ inches against a spring, in order to release any object which might be caught. Power cannot be applied until all side doors are closed. Provision has been made to by-pass the mercury switch which gives this protection, at any door, in any car, or on the whole train.

The doors are pneumatically operated with an air engine for each section, 12 per car. A rotating engine arm with an overcentre device locks each section in the open or in the closed position. Two side doors of each car may be opened by passengers in an emergency, either with or without the aid of car air pressure.

HEATING AND VENTILATION

The cars are heated by thermostatically controlled air passing over electric heaters. A 30-kw electric heater, mounted underneath the car in an insulated "oven" provides the heat for the air forced through it by a motor-driven fan, thence through a system of insulated ducts to louvres under every seat through which it is blown at low velocity. The system is designed to blow 2,000 cubic feet of air into the car per minute. Of this amount, 50 per cent is fresh air drawn in from the outside and 50 per cent is exhausted air from the car through two rows of louvres at the edge of the ceiling.

The car temperature is maintained at approximately 65 degrees Fahrenheit, and the heating elements are in three circuits of 10 kw each. During preheating before entering service, or at such times as the entire 30 kw is required, the fresh air sup-

Fig. 1. Eglinton Terminal, the northerly terminus of the Yonge Street Subway, showing surface bays used by various connecting trolley coach and bus routes

Fig. 2. Eglinton Terminal. Platform level showing terrazzo flooring, glass-faced masonry, acoustic ceiling, tubular station lighting, etc. Center platforms are used at terminals only

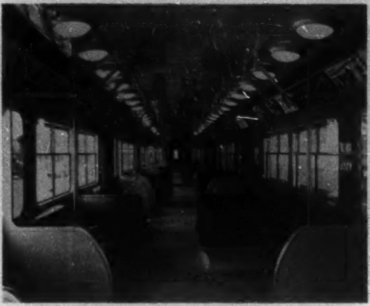


Fig. 3 (left). Interior view of rapid transit car showing seating arrangement, heating ducts, ceiling-mounted standee stanchions, etc.

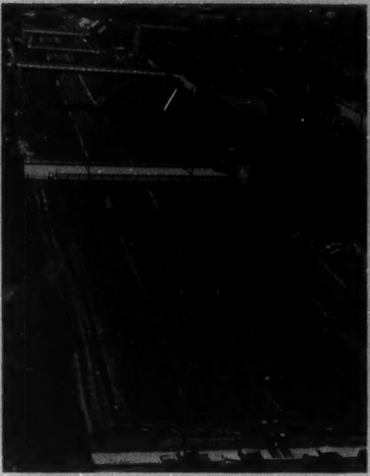


Fig. 4 (right). Open-cut section of the subway looking south from St. Clair showing clean lines of bridges and side slopes

ply is temporarily shut off. In the warm season, the heaters are rendered inoperative and the system provides fresh outside air only. Proper safeguards have been provided to prevent overheating of the heating elements in case of failure anywhere in the system.

LIGHTING

The interior car lighting is provided by 47 lights arranged in two rows of ceiling-mounted incandescent fixtures spaced $2^1/_2$ feet apart with a 48-watt 30-volt bulb per fixture. The fixtures are circular with an attractive anodized aluminum frame, and have been designed to provide good general lighting, plus an abundance of light on the reading plane, about 20 foot-candle light. The body lights in each car are under control of a switch in the cab. They are lit at all times when the cars are in service as the cars run in light and darkness alternately in rapid succession.

Emergency body lighting is provided by operating six of the body lights, one over each doorway, from the battery. A 2-pole switch in each cab controls the lights, one pole controlling the 600-volt circuit and the other one the battery circuit. This permits elimination of a relay to energize the battery lights in case of a power failure. Miscellaneous lighting includes marker lights, destination signs, run number, external door signal lights, and cab signal lights.

MOTORS AND CONTROL

All axles are motored, four per car, each motor of 68 horsepower, and connected two-in-series. They are mounted longitudinally on the trucks. They transmit power to their respective axles by a propellor shaft and hypoid gear drive. Each gear unit is pressed directly on a forged steel axle and is prevented from turning by a torque arm connecting it to the truck frame. The motors are connected in series-parallel, with three speeds, "switching," "series," and "parallel." Balancing speed of an empty train on level tangent track with 600-volt supply is 50 mph, which is attained with the motor fields shunted.

Control of the motors is by pneumatic cam magnetic control. It has several unique features not common on this Continent. The master controller is of the desk type, the bottom portion located separately in order to save cab space by permitting the operator to sit closer to it with his knees underneath it. The remaining controller functions, other than to control motion and direction, are located in another near-by control cabinet. The usual deadman feature of the controller is nullified when the reverser handle is in the neutral position, thereby permitting yard parking without causing an emergency brake application.

The acceleration is automatic, with a choice of 3 rates under control of a switch which allows a maximum rate of 2.3 miles per hour per second (mphps) on level track under most conditions of loading. A sequence light in the motorman's cab is illuminated at the end of each acceleration after every motor on the train has reached its shunted field position. The control can be similarly checked in a few seconds with the train stationary and, in case of a defective car, test facilities permit it to be localized quickly. These two features have proved very valuable.

Low-voltage energy is supplied by a 50-volt 4-kw motor-generator set and a 23-cell battery located on each *A*-car. The energy thus produced is sufficient for the two cars of each unit while the compressed air, generated on the *B*-car likewise supplies both cars.

THE CAB

The cab is an excellent example of flush mounting equipment which does not materially interfere with its accessibility. Much of the electrical equipment is mounted on the inner side of hinged panels which swing out for inspection or which can be disconnected, unhinged, and taken to the bench for major work.

The duties of the guard are performed in a cab not occupied by the motorman. The full-drop counterweighted cab side-window gives the guard a quick, easy, and complete view of the platform. Control of the doors is by push buttons sending momentary opening or closing impulses to the door engines. The establishing of a guard's position in any cab is extremely simple, requiring only the turning of a key and the moving of a handle. If more than one guard is on duty, the door control is divided automatically into the same number of zones as guards, by the simple action just described. Failure to establish the proper start, end, or operating position of a zone is impossible.

PNEUMATIC SYSTEM

Compressed air is used to stop the train and to operate the doors and several parts of the control. It is obtained from a 43 cubic feet per minute lightweight air compressor with a built-in alcohol anti-

Fig. 5. Exterior view of a 2-car train taken at the entrance to the repair shops at Davisville

Fig. 6. View of Wellesley Station showing modern design of surface station buildings. Surface vehicles loop around this building for passenger interchange

freezing dispenser in the intake. All the compressed air is passed through a filter unit from which it emerges in three circuits for brakes, doors, and control. For each of the door and control circuits the air is also reduced in pressure, gauged, and provided with cutoff facilities as well as with a mist lubricator for the internal moving parts of the equipment served. When the air valve to the door engines is closed, it automatically closes the doors and by-passes the line-switch mercury door-interlocks of all doors on the same car at the same time.

The brakes are of the standard electro-pneumatic self-lapping type but are unique in having one brake cylinder per shoe, or 16 per car, thereby eliminating rod and levers, and noise. Each cylinder has an automatic slack adjuster and a method for altering the leverage ratio. In case of failure of the electric portion of the system, the standard automatic brake which operates by variation of train pipe pressure is available through further movement of the brake valve handle.

A retardation controller has been provided to permit rapid, smooth stops with minimum attention from the operator. It consists essentially of two mercury U-tubes, each set to close an electric contact at given rates of braking. One tube prevents further admission of air to the cylinders when a braking rate of 2.8 mphps has been reached; the second releases air from the cylinders when the braking rate reaches 3.2 mphps. Thus, under ordinary conditions, the operator need only move the brake valve to the full application position and do nothing more until the train is almost stopped, when the pressure should be gradually reduced to avoid a lurch. Proper condition of the retardation controller is indicated by a pilot light in the cab, and functioning of the unit during a brake application is indicated by another cab pilot light going on and off.

Emergency application of the brakes may be executed by the operator; by the guard from any cab; or by a passenger at one location in a car. Emergency brake application may also be made at any of the usual points in the train where the brake pipe may be vented, or at the trip valves which are carried at the two diagonal corners of each 2-car unit.

The trip valves are installed to perform the usual function of forcing the train to stop if it should pass the raised trip arm of a wayside stop signal. The trip valves have several unusual features. On a train, only the end trip valves function, the intermediate ones being rendered inoperative automatically during coupling, while they are restored to service automatically during any uncoupling, which places them again at the end of a train. Inoperative trip valve arms will, if struck by a track trip, latch up in a raised position so that they will not be hit twice, but will be restored automatically to their vertical operating position whenever an uncoupling operation places them at a train end. Furthermore, the arms are trailable, i.e., if a train were operating in the reverse direction on a track and the trip arm valve were hit by the track trip in the opposite direction to normal, the brakes would not be applied.

Protection against certain abnormal conditions has been provided by a safety valve to prevent excess air pressure in the brake cylinders and by emergency means for releasing the cylinder pressure on either truck from within the car or on both trucks from underneath it.

The car handbrake is applied on only the cab end truck of each car. It is of the ratchet type, and the operating mechanism is housed in a covered recess outside the end door.

Coupling Equipment

The connections between cars are of two kinds. Between the two cars of each unit there is a mechanical bar connector and a spring buffer for the mechanical forces, with hoses and multiconductor cables for the pneumatic and electrical connections respectively. Each end of the multiconductor cable is joined to the adjoining car by a plug and receptacle with, however, the added feature of a pin whose shearing upon the breaking apart of two cars will enable the electrical connections to separate without damage.

The connection between 2-car units is accomplished by an automatic coupler, cab-controlled by push buttons. The spring bumper also assists in minimizing relative movement between adjacent cars and in providing a smooth walking surface between cars. The automatic coupler is of the wedgelock type in which the connecting tongues are held in contact by spring-applied air-released wedges. In uncoupling, the spring bumper shoves the cars apart a few inches upon disengagement of the coupler tongues.

Shipment

Shipment of the cars from Gloucester to Toronto presented unusual problems. Since the gauge is 4 feet $10^{7}/_{8}$ inches, and not standard, the car bodies were hauled from the factory to the seaport at Bristol on special factory trucks while the subway trucks were shipped on flatcars. Since the couplers were not the same as those on the railroad, a short freight car was equipped with a railroad coupler on one end and a subway coupler on the other end.

Since the cars are 10 feet 4 inches wide, their roof eaves would have fouled some of the bridges under which the tracks ran; therefore, when it was necessary to pass such obstructions, a special feature of the factory trucks permitted each car to be cranked sideways by 8 inches on either side of its center line.

After crossing the Atlantic, with the bodies as deck cargo and the trucks in the hold, the vessels docked at Montreal during the navigation season, and at Halifax, Nova Scotia, or Saint John, New Brunswick, when the St. Lawrence River was closed to shipping. Special tackle was designed and forwarded to each port to permit the huge floating cranes to load and unload the cargo. At the Canadian ports, railroad flatcars equipped with rails at Toronto gauge were placed so that the floating cranes could transfer the trucks to the rails and the bodies to the trucks. The rail haul to Toronto completed the journey, all of which was accomplished with negligible mishaps.

Conclusion

The Toronto subway system has demonstrated its ability to restore the downtown area to its rightful position as the heart of the city. It has slowed down the flight to the suburbs. It has provided a positive rebuttal to those who urge that the provision of more facilities for the motor car is the only answer to traffic ills. It has proved that a vigorous transportation agency, alert to the varying needs of the area it serves, can face the future with confidence. And it provides a challenge to AIEE members to gather and provide facts and information so that other transportation systems may become leaders rather than followers in their communities.

◆

Discussion

C. M. Hines (Westinghouse Air Brake Company, Wilmerding, Pa.): The present paper reveals that the Toronto Subway is now an accomplished fact. It is perhaps due to the modesty of the author that this paper is limited to factual statements and omits any mention of the enormous difficulties which had to be overcome in completing this project.

I was present in Toronto on March 30, 1954, when the subway was opened. It was readily apparent that the people of Toronto were justifiably proud of the accomplishments of the Toronto Transit Commission.

One apparently small detail was perhaps more truly significant of the manner in which the Toronto Transit Commission worked than any other. I had heard Mr. Inglis state on several occasions that, when the subway was operating, the surface cars would be removed from Yonge Street. The dedication ceremony on opening day began about 11:30 a.m. The subway was opened to the public at about 1:00 p.m. The last surface car was removed from Yonge Street at 4:00 p.m. of the same day and work was begun at 7:00 a.m. the next morning to tear up the tracks in the street.

Although the rolling stock was purchased in England, we were fortunate in being allowed by the Toronto Transit Commission to conduct some experiments on a new type of composition brake shoe being developed jointly by ourselves and the Johns-Manville Corporation. One of the many advantages of such a brake shoe is the removal of the source of cast iron dust with its effect on the insulation of electric devices.

Other equally important advantages are long brake shoe life, better wheel condition, smoother stops, and quieter operation. Tests on this form of brake shoe are continuing on an increasing scale and the results so far have been encouraging.

Overvoltage Detection in Single and Multigenerator Aircraft A-C Systems

W. M. TUCKER
MEMBER AIEE

M. TRBOVICH
MEMBER AIEE

OVERVOLTAGES resulting from fault conditions in aircraft a-c electric systems being designed at present can cause extensive utilization equipment damage. Until such time as aircraft power generating and regulating equipment design results in reduced overvoltage possibilities, it is likely that an overvoltage control relay will be installed in many single-generator systems.

Generators operating in parallel have nearly the same terminal voltage, since they are connected to the same bus by cables having a small voltage drop. If a fault develops, such as a break in a voltage-sensing lead to the regulator, the associated generator normally attempts to deliver overvoltage. A large circulating current flows between generators, which nearly equalizes the terminal voltages at some intermediate value. The more generators in parallel, the nearer normal the bus voltage remains. With many generators in parallel, the bus voltage might not rise sufficiently to operate the overvoltage control relay, but both normal and faulty generators could be damaged by the large circulating current and by the high field current, if allowed to exist for a long interval. With fewer generators, the terminal voltages rise until all overvoltage control relays operate, thus destroying the continuity of power supply. Therefore, in a paralleled generator installation, the addition of a selective overvoltage control relay will increase the electric power system reliability in terms of continuity of regulated electric power.

To insure continuity of power under the fault conditions described, a selective voltage detection circuit is proposed. This circuit provides a positive bias voltage signal which, when added to the generator terminal voltage, enables the overvoltage control relay to operate and remove only the faulty generator from the bus. The other overvoltage relays (OVR) in the system will receive a negative bias voltage signal to keep them from operating. However, the effectiveness of an overvoltage control relay depends upon an adequate description of its required operating characteristic, for false operation is less desirable than no overvoltage control.

Further, reference 1 describes the characteristics of military aircraft electric power. This document places a limit on the line-to-neutral voltage which may exist in the system from any cause. At present, this voltage limitation must be accomplished by an OVR. The OVR

Paper 54-327, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted April 22, 1954; made available for printing June 23, 1954.

W. M. Tucker and M. Trbovich are with the Naval Research Laboratory, Washington, D. C.

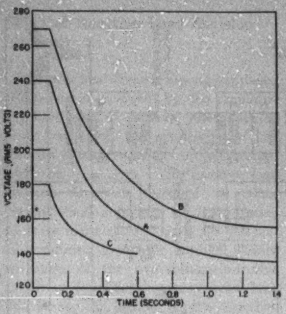


Fig. 1. Proposed overvoltage control relay characteristic

Curve A—Maximum voltage versus time during which the relay must not operate, the lower limit of the tolerance band
Curve B—Proposed upper limit of the tolerance band; the proposed maximum voltage versus time at load equipment
Curve C—Existing overvoltage curve in Military Specification MIL-E-7894,[1] Fig. 1

characteristic therefore influences the maximum voltage, which will be applied to the terminals of utilization equipment.

OVR Minimum Operating Characteristics

In view of the large quantity of data available on power system characteristics and the lack of quantitative information regarding overvoltage limitations of utilization equipment, it is expedient first to describe overvoltage control in terms of maximum voltage versus time which must not energize the OVR.

It is characteristic of aircraft 3-phase a-c generator-regulator units that an overcurrent on one phase may cause an overcurrent and an overvoltage to exist simultaneously in the system. The purpose of a circuit protector is to isolate faulted circuits and thus permit the remainder of the system to continue normal operation. The function of over voltage protection in single-generator electric systems is to prevent damage to utilization equipment by de-energizing the generator when system overvoltage exists. In multigenerator electric systems, the overvoltage protection circuit has the additional task of protecting the normally operating generators. In view of the preceding component functions, it follows that a circuit protector should isolate distribution system faults before the OVR is energized by the overvoltage on the unfaulted phases.

Table III, reference 2, shows the permissible circuit protector rating for a fixed fault duration. Fig. 10, reference 2, shows the resulting unfaulted phase voltage corresponding to the same fault durations. With these data at hand, it is necessary to select reasonable values of circuit protector ratings and overvoltage. Curve *B*, Fig. 10, reference 2, which has been extended and reproduced as curve *A*, Fig. 1, of this paper was selected to represent the largest line-to-neutral voltage resulting from line-to-neutral faults which may be allowed to exist at the terminals of utilization equipment and which should not energize the overvoltage control relay. Practically, this curve describes the minimum operating characteristic to be allowed to operate the overvoltage control circuit. Based upon power system data on hand, curve *A*, Fig. 1, is a realistic OVR characteristic for application in a single-generator installation with either differential-current type of feeder-fault protection or with no consideration given to feeder faults.

A relay characteristic, however, is described in terms of an operating band. Also, multigenerator installations are capable of performing either as a paralleled system or as isolated systems. Therefore, the circuit protection components must also be designed to perform in both types of systems which, for the OVR, means that the maximum operation time should also be specified. In determining permissible operation tolerances for the overvoltage control relay, two primary factors must be considered: 1. maximum operation tolerance permitted while retaining the selectivity in parallel generator operation, which is considered in this paper; 2. reasonable production tolerances.

Selective Circuits for Parallel Systems

Proposed Circuit

The selectivity feature of the overvoltage control relay proposed for paralleled system operation requires bias voltage signals which are added vectorially to the voltage normally sensed by the overvoltage control relay. Use is made of the fact that when the field excitation of a generator operating in parallel with others is raised, its load current becomes more lagging while those of the other paralleled generators become less lagging or possibly leading. This is owing to a reactive circulating component of current between generators. No real circulating component exists unless caused by the prime mover load characteristics.

Fig. 2 shows a circuit in which bias voltage signals of the proper phase relation provide a higher voltage signal from the generator causing overvoltage and a lower voltage signal from the remaining generators. In this figure, two generators are shown, but the selective circuit may be used with any number of generators. The proposed bias voltage circuit is shown for one phase.

In Fig. 2, a current transformer (CT) with secondary terminals, *X*1 and *X*2, is placed in phase *A* of each generator.

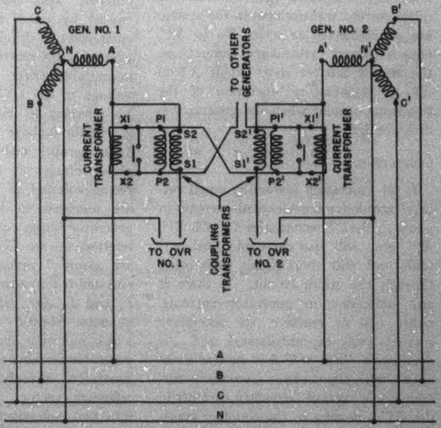
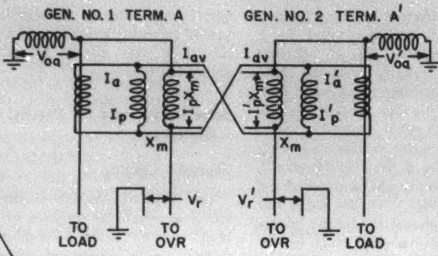
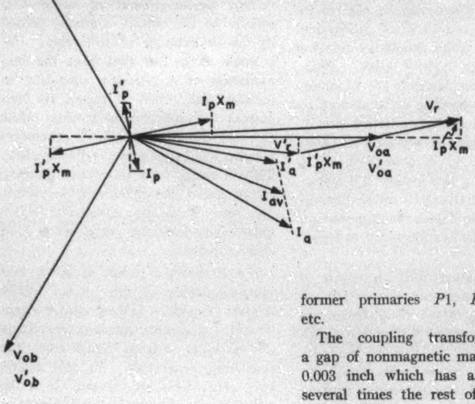


Fig. 2. Overvoltage selectivity circuit for use with line-to-neutral sensing OVR





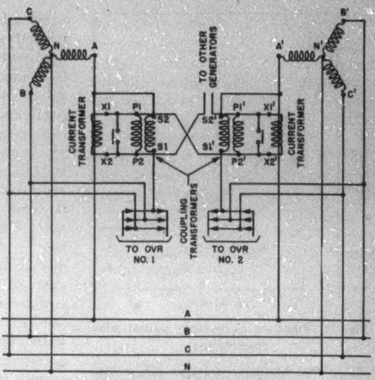


Fig. 3 (left). Vector diagram of voltages and currents in overvoltage selectivity circuit

Fig. 4 (above). Overvoltage selectivity circuit for use with 3-phase average-voltage sensing overvoltage relay

The primary winding of a coupling transformer, $P1$ and $P2$, is connected to terminals $X1$ and $X2$. The secondary winding of the coupling transformer, $S1$ and $S2$, is connected to the generator terminal A and in series with the overvoltage relay coil. Terminals $X1$, $X2$, $X1'$, and $X2'$ and like terminals of other current transformers in the system are all connected in series $X2$ to $X1'$, etc., as shown. Auxiliary contacts on the circuit breaker short-circuit $X1$ to $X2$ for satisfactory operation of the remaining system when the generator is disconnected.

OPERATION

If all generators (assuming equal ratings) are delivering identical currents in phase A, the CT secondaries $X1$, $X2$, $X1'$, $X2'$, etc., will also carry identical currents and will short-circuit each other through the series circuit. If there is any unbalance in generator currents, either real or reactive, the secondary currents will be unbalanced and the average of these will flow in the series circuit with the excesses (secondary current minus average secondary currents in the CT's) flowing in the coupling transformer primaries $P1$, $P2$, $P1'$, $P2'$, etc.

The coupling transformer contains a gap of nonmagnetic material of about 0.003 inch which has a reluctance of several times the rest of the magnetic circuit. Since the permeability of this gap is constant, a nearly constant ratio of magnetic flux to primary current exists, and therefore secondary induced voltage is nearly proportional to primary current. The bias signal voltage is therefore nearly proportional to the circulating component of the generator current. This bias signal voltage is caused by either a real or a reactive component of circulating current. However, only the reactive component has an appreciable effect on the total signal voltage to the OVR. This can be seen by referring to Fig. 3. Let V_{oa}, V_{ob}, V_{oc} be the 3-phase voltages of generator 1 in parallel with generator 2 whose phase voltages are V_{oa}', V_{ob}', V_{oc}'. The phase voltages of the two generators are assumed to be coincident. The generator phase currents are I_a and I_a', referred to the CT secondaries, which are assumed unbalanced, having both real and reactive circulating components. I_p and I_p' are generator current components referred to CT secondaries, which represent the differences between generator currents and the average of the generator currents. Let V_{oa} be the reference vector. I_p can be resolved into two components, one reactive and one real where the reactive component results from field excitation unbalance, and the real component results from power input unbalance.

I_p flows through the coupling transformer primary and induces the secondary voltage I_pX_m. X_m is the mutual reactance between primary and secondary. I_pX_m is the bias signal voltage and may be resolved into two components. The real component results from the reactive component of I_p and therefore from the field excitation unbalance. It adds directly to V_{oa}. The reactive component of I_pX_m, which is the result of the real component of I_p and therefore of the power input unbalance, is at right angles to V_{oa} and, if small compared to V_{oa}, has little effect on the resultant magnitude.

The resultant voltage V_r is the signal to the OVR of generator 1. In a similar manner V_r' is the signal to the OVR of generator 2.

Fig. 4 shows a circuit similar to Fig. 2 except that the OVR is connected to the output of a full-wave 3-phase rectifier connected to the generator terminals, and senses the approximate average of the 3-phase voltages, one of which is modified by the bias signal voltage.

It is not necessary that all generators in parallel be of the same rating for the proper operation of this circuit. The CT's should have secondary to primary turns ratio proportional to generator ratings so that the CT secondary currents would be equal when rated currents flow through the CT primaries.

Effect of Reactive Load Division Circuit

Reactive load division among paralleled generators is governed by the regulated characteristics of each of the paralleled units and upon the sensitivity of the reactive load division circuits associated with the regulators. The reactive load division circuit is designed to reduce reactive load unbalance so that system capacity is maintained above 90 per cent of combined generator ratings. Since operation of the overvoltage selective circuit is dependent upon reactive load unbalance, the action of the reactive load division circuit is directly opposite to that desired for overvoltage selectivity. No difficulty exists, however, since sufficient reactive load unbalance is allowed to exist to obtain the required bias voltage signal.

Sensitivity Requirements

An accurate overvoltage detection circuit would be one in which the bias voltage signal is small so that the OVR senses the actual bus voltage. This is neither practical nor reliable. First, it has been stated that when a generator is operating in parallel with others it does not produce as high an overvoltage as when operating alone; the more generators in parallel, the less is the overvoltage. If the bias voltage is too small, the OVR will not operate and the generator may be damaged by overheating from the circulating currents. On the other hand, if the bias voltage signal is too large, the resultant voltage to the relay will be undesirably influenced by real load unbalance.

The sensitivity of the overvoltage selective circuit may be defined as the change in the voltage signal to the OVR operating coil per ampere difference from average of generator reactive current expressed in per unit. The base current is rated load current, and not rated reactive load current. With these definitions, the sensitivity will not depend on the method of sensing such as line-to-neutral or 3-phase averaging, nor on the rating of the generator. The bias voltage signal, however, must be greater for 3-phase averaging if only one bias voltage signal per generator is used.

The sensitivity proposed for the overvoltage selective circuit based on test data and the other considerations mentioned is from 0.25 to 0.30 volt-per-ampere reactive circulating current (for 30-kva generators) or an average of 0.19 per unit for all generator ratings. This figure is based on the existing reactive load division circuit sensitivity as found for the test circuit. The sensitivity of the bias voltage circuit and the overvoltage tolerance band are related, in the sense that increased sensitivity permits a wider tolerance band.

Test Circuit and Experimental Data

A series of tests of the proposed circuit (Fig. 2) was made to determine the reliability for selective tripping and absence of nuisance tripping. Two OVR's used were adjusted as closely as possible for identical voltage-time characteristics. Two 30-kva narrow speed-range generators with magnetic amplifier regulators were operated in parallel. Tests included overvoltages caused by an open voltage-sensing lead to the regulator both in and not in the phase supplying power to the magnetic amplifier under various load conditions, and were made with and without use of the reactive load division circuit. Included in the tests was the effect of opening a field circuit at no load, both with and without the reactive load division circuit. With no system load, and the field circuit of one generator open-circuited, the unfaulted generator reaches ceiling excitation and the bus voltage drops. When load is connected to the system, the bus voltage decreases further. OVR nuisance tripping does not occur in spite of a larger bias voltage signal. A theoretical study shows that if the generators lose synchronism when the field circuit of one generator is open circuited, the voltage signal to the OVR will not be sufficient to cause nuisance tripping.

Table I. Test Data on Selective Overvoltage Circuit under Fault Conditions

Test Conditions	Type of Fault	V_{BUS}, Volts	I_{G1}, Amperes	I_{G2}, Amperes	I_L, Amperes	OVR_1, Volts	OVR_2, Volts	OVR Diff., Volts	Frequency, Cycles
A1	C	186	85 /90	85 /−90	0	165	208	43	380
A1	C	190	103/−90	96 /90	0	215	168	47	380
B1	C	185	88 /90	88 /−90	0	164	206	42	380
B1	C	190.5	105/−90	95 /90	0	215	168	47	380
A3	C	158	45 /90	151 /−90	113 /−90	140	180	40	380
B3	C	155.5	57 /90	162 /−90	110 /−90	135	179	44	380
A7	C	136	41 /90	190 /−90	165 /−90	142	162	20	380
A7	C	130.5	210/−90	54 /90	163 /−90	163	115	48	380
B7	C	126.5	52 /90	196 /−90	155 /−90	133	160	27	380
A1	D	144.5	45 /90	46 /−90	0	131	157	26	380
A1	D	141	50/−90	44 /90	0	151	130	21	380
B1	D	150	48 /90	50 /−90	0	136	163	27	380
B1	D	145.5	48/−90	43 /90	0	156	134	22	380
A2	D	131.5	16 /90	106 /−90	94 /−90	114	149	35	380
B2	D	132.5	33 /90	122 /−90	94 /−90	112	154	42	380
A3	D	132.5	17 /90	107 /−90	94 /−90	116	149	33	380
B3	D	134.5	40 /90	131 /−90	96 /−90	112	158	46	380
A4	D	129.5	84 /71	91 /−61	72 /0	104	155	51	380
B4	D	135	91 /78	104 /−59	73 /0	104	164	60	380
A5	D	126.5	94 /71	96.5/−67	69 /0	104	153	49	380
B5	D	134.5	92 /74	102 /−61	74.5/0	107	165	58	380
A6	D	130	26/−90	133 /−90	159 /−90	116	144	28	380
B6	D	125.5	25 /90	167 /−90	159 /−90	104	151	47	380
A7	D	129.5	30/−90	131 /−90	159 /−90	116	143	27	380
B7	D	126.5	30 /90	167 /−90	155 /−90	107	151	44	380
A8	D	150.5	45 /90	48 /−90	0	136	163	27	380
A8	D	158	37 /90	41 /−90	0	145	168	23	420
A8	D	126.5	8 /90	11 /−90	0	122	128	6	420
A8	D	124	11/−90	11 /90	0	120	114	6	380
A1	E	112	55/−90	51 /90	0	128	97	31	420
A1	E	113	41 /90	43 /−90	0	99	126	27	420
B1	E	116	62/−90	57 /90	0	133	98	35	420
B1	E	117	42 /90	44 /−90	0	104	130	26	420

Test Conditions

A—reactive load division circuit in use
B—reactive load division circuit not in use
1—no load
2—83.4-ampere load current, 0% power factor, (pf) lagging, equally divided
3—83.4-ampere load current, 0% pf lagging, divided 50 to 40
4—62.6-ampere load current, 100% pf, equally divided
5—62.6-ampere load current, 100% pf, divided 50 to 40
6—144-ampere load current, 0% pf lagging, equally divided
7—144-ampere load current, 0% pf lagging, divided 50 to 40
8—no load, regulator 1 adjusted for minimum voltage, regulator 2 adjusted for higher voltage to cause a circulating current of 30 to 34 amperes, the highest value which will not cause OVR to trip before establishing fault.

Type of Fault

C—open sensing lead in phase A to the regulator of generator supplying equal or lesser load
D—same as C except in phase B
E—open exciter field circuit

Data obtained from these tests are shown in Table I. It is noted that proper selective tripping by the OVR occurred in every case where desired and no-nuisance tripping occurred. The values of generator and load currents listed are approximate, especially with regard to phase angles, since it is difficult to obtain a purely reactive load and to have no real load unbalance at all times. With two paralleled generators operating normally, an open loop in the series circuit connecting CT secondaries does not cause a nuisance trip unless the reactive load of either generator is increased to 1.25 per unit.

The sensitivity of the bias voltage circuit was approximately 0.25 to 0.30 volt-per-ampere circulating current. A center tap was installed on the coupling transformer secondary to vary this sensitivity if desirable.

Overvoltage Control Relay

Operation Tolerance

The effectiveness of the proposed selectivity circuit is dependent upon the operation tolerance of the basic overvoltage control relay. For reasons given earlier, Fig. 1, curve *A*, is considered to be a reasonable minimum characteristic for this relay.

Measurements on the proposed selectivity circuit with the recommended sensitivities also given earlier in this paper show that for the various fault conditions, under which the OVR is required to operate, the difference in applied voltage to the two units ranges from 20 volts for low-current faults to 30 volts for high-current faults.

The data in the two preceding paragraphs permit the description of a characteristic band which the basic overvoltage control relay must operate (Fig. 1).

Maximum System Voltage

Reference 1 describes the maximum line-to-neutral voltage which may exist at the terminals of utilization equipment but which should not damage this equipment. This document describes the maximum line-to-neutral voltage which may exist in the system from any cause.

An extensive study of unfaulted phase overvoltage during faults was made on generators and regulators available at the Naval Research Laboratory.[2] These observed voltages exceeded the values given in reference 1. To meet the existing overvoltage requirements of reference 1: 1. many of the generators and regulators must be redesigned; 2. the distribution system must be redesigned to require the low circuit protector ratings only (below 5 amperes); or 3. a fast-acting distribution-system-wiring protector must be developed. Since reference 1 should depend upon the electric power characteristics in present and immediate-future aircraft, a revision in the specified overvoltage requirement is necessary.

It is assumed that Fig. 1 represents an operating band to which an OVR of reasonable weight and size can be designed. If this assumption is correct, curve *B*, Fig. 1, represents the maximum line-to-neutral voltage which would exist on the system for any reason, and therefore describes the maximum line-to-neutral voltage which may exist at the terminals of utilization equipment connected to electric power systems employing a variety of generating and regulating equipment.

Conclusions

1. Selective overvoltage detection of generators operating in parallel can be obtained reliably and directly by making use of bias voltage signals obtained from the circulating currents between generators.

2. The sensitivity of the bias voltage circuit is dependent upon the sensitivity of the reactive load division circuit and on the tolerances required by the overvoltage relays.

3. In the proposed system using typical components such as have been available at the Naval Research Laboratory, the maximum line-to-neutral voltage which will exist at the terminals of utilization equipment is described by curve *B* of Fig. 1. This conclusion is based upon the assumption that the operation band of an OVR as described in Fig. 1 can be reasonably manufactured.

References

1. Characteristics of Aircraft Electric Power. *Military Specification MIL-E-7894*, Bureau of Aeronautics, Washington, D. C., April 28, 1952.

2. Co-ordination of Overvoltage and Short-Circuit Protection for Aircraft A-C Systems, **M. Trbovich, R. E. Kidwell, Jr., W. M. Tucker.** *AIEE Transactions*, vol. 72, pt. II, 1953 (Jan. 1954 section), pp. 369–74.

◆

Discussion

R. W. Stineman (Boeing Airplane Company, Seattle, Wash.): Curve *A* in Fig. 1 is represented as the maximum voltage which must be sustained by the loads in an aircraft a-c system. The point that this curve is well above presently specified voltage limits is well taken. However, the authors' immediate conclusion that curve *A* must be the minimum operating voltage for the OVR is not well founded. The basic sources of overvoltage may be classified as follows:

1. Faulty voltage regulating system.
2. High voltage on unfaulted phases during an unsymmetrical short circuit.
3. Transient voltage overshoot following removal of a load or short circuit.

The OVR should operate for category 1, but should not operate for categories 2 and 3. If the OVR has the same type of sensing as the voltage regulator, e.g., average line-to-line voltage sensing, then the OVR properly performs its intended function for category 1 and has no cognizance whatever of the overvoltage in category 2. The only logical purpose for the time delay incorporated in the OVR is to permit it to override voltages in category 3. Going back to the origin of curve *A* (ref. 2 of the paper), we find that it is a composite of categories 2 and 3. Thus, curve *A* represents the voltage which must be sustained by load equipment, but which has no direct relation to the OVR.

It may be noted that protection for the loads against category 2 over-voltage is provided by fault detection and isolation. The OVR has no relation to this condition other than that it should not operate.

In a previous paper[1] I describe another means of providing overvoltage protection in a paralleled a-c system. The philosophy of the scheme is that the existence of large differences among the current outputs of several alternators is an indication of improper paralleling and requires that an overexcited alternator be isolated from the other alternators but not disconnected from its loads. Any alternator which is isolated is provided with conventional overvoltage protection. The differential reactive-current protection relay has the following advantages over the current-biased OVR proposed by Mr. Tucker and Mr. Trbovich:

1. There is no possibility of falsely tripping off any alternator during such conditions as loss of one alternator's excitation, a broken equalizer loop, or bus faults on the equalizer phase.

2. Protection is provided at no additional cost against a broken equalizer loop, a bus fault on the equalizer phase, 3-phase faults, an alternator failure, or an unstable voltage regulator; all in addition to overvoltage protection.

I am not aware of any disadvantages of the differential reactive-current protection relay compared with the current-biased OVR. Perhaps the authors can offer something along this line.

Reference

1. Differential Reactive Current Protection Relay, **Russel W. Stineman.** *AIEE Transactions*, vol. 72, pt II, 1953 (Jan. 1954 section), pp. 407–12.

T. B. Owen (Douglas Aircraft Company, Santa Monica, Calif.): I have been following with great interest the series of Naval Research Laboratory reports which have dealt with the general subject of overvoltage and its detection in aircraft. The present paper depends heavily upon generator overvoltages determined previously, and some serious questions as to the interpretation of data must be raised. The conclusions given in the present paper involve such a radical change of current specification requirements (see ref. 1 of the paper) that the supporting data must be critically reexamined.

In Fig. 1 the present specification overvoltage requirements are compared with proposed requirements. These proposals are based on tests reported in reference 2 of the paper, which gives in some detail data on overvoltages encountered under test conditions with a series of aircraft generators. It is these data that I wish to examine.

In reference 2 of the paper six generators, generators A through F, are described. Further data on these machines are:

Generator A. X_d is 2.0 per unit, no-load saturation voltage the highest yet encountered (160 volts at 1.75-per-unit excitation, curve rising very sharply but undefined past this point). Not in production, only a few made, test machines only.

Generator B. No longer listed in the manufacturer's catalogue. Low reactance and no-load saturation voltage.

Generator C. This is the A-1 machine; it is not being used in current design. Large number in use in older installations. Low reactance and no-load saturation voltage.

Generator D. Current production machine. Low reactance, maximum no-load saturation voltage of 185 volts.

Generators E and F. Two samples of the same model. There are apparently two or more versions of this machine; one is listed by the manufacturer as having an X_d of 2.65, another as having an X_d of 2.00. The test on this machine available to me shows an X_d of 2.35 and a ceiling no-load voltage of 210 volts.

It seems quite clear that generators A and B are not representative of aircraft machines, and should not be used as examples. Generator C should also be deleted; the machine which replaces it has somewhat different characteristics. With these three generators eliminated from consideration, only two are left, with rather widely differing characteristics. In any event, using these machines, neither curve *A* nor *B* of Fig. 1 is justified. The maximum overvoltages are closer to curve *C*, which is the present specification requirement (ref. 1 of the paper).

An analysis of the data presented in Fig. 5 shows that the only time curve *C* of Fig. 1 is exceeded is when the generators are operating at no load and are at ceiling excitation. This is an unreal requirement, inasmuch as even a small amount of load brings the phase voltages down considerably.

The whole point of this discussion is that you must not specify system parameters upon a component basis. Overvoltage curves must not be constructed on the basis of what the generators will supply, but on the basis of what the using equipment will stand when such equipment is constructed according to the best standards available. It seems to me that a rather general tendency is to think of the generators as being the most important part of the electric system, and to think of protection in terms of the generator. After all, particularly in military airplanes, the only reason in many cases for the existence of the airplane is to take a load of electric or electronic equipment to a place where it can be used to best advantage. The generator exists to serve this equipment; an overvoltage protective scheme which burns out or causes this equipment to malfunction is added useless weight and is a hazard. The object of overvoltage protection is to protect the using equipment; if, in the process, the generator can also be saved, all well and good, but the prime consideration is that the generator be taken out of service before the equipment is damaged or the mission is aborted.

I believe that overvoltage curves must be specified on two bases:

1. What can presently available equipment withstand without failure or malfunction, and what would be the weight and time penalties incurred by a redesign to higher standards? An illustration of this is the question of circuit breakers: in many instances present *AN3160* and *AN3161* breakers are inadequate where the generator recovery voltage is too high, but adequate where the recovery voltage is low. Which is preferred, redesign of circuit breakers or a restriction on recovery voltages?

2. Can overvoltage selective circuits be designed which will take into account the considerations of basis 1, the obtainable relay tolerances, and reactive load division circuit constants, and give a system which will select the faulted machine and isolate from the other machines with a minimum of disturbance? This requirement is analyzed in this paper; the conclusion that such a requirement is obtainable is certainly justified.

In conclusion, I must say that I am very much indebted to Mr. Tucker and Mr. Trbovich for their analysis of the circuitry needed to protect against the overvoltage condition. I merely wish to protest against their generalizing from inadequate data. It is a serious problem, and it must be attacked from the proper angle. We cannot afford to tie future designs to obsolete equipment.

W. M. Tucker and M. Trbovich: We appreciate Mr. Stineman's and Mr. Owens' interest in our paper and the latter's interest in our work of the past few years.

Mr. Stineman questions the functional relation between overvoltage protection and overvoltage from unbalanced faults. Many of the electric system loads in aircraft are single-phase and line-to-neutral connected. For this reason Military Specification *MIL-E-7894* specifies that the line-to-neutral voltage remains within a certain range of magnitudes. It is important that voltage outside this range be isolated from utilization equipment. An OVR characteristic must therefore be related to line-to-neutral voltage regardless of the manner in which overvoltage is sensed.

We note that the caption of Fig. 1 is not as clear a statement of intended interpretation of this figure as is the text. We identify the curves of Fig. 1 as the maximum line-to-neutral voltage which should not operate the OVR. The applicability of this statement is independent of the overvoltage sensing method used. Fig. 1 is, indeed, related to the OVR characteristic and the manner in which it is related depends upon the sensing circuit of the OVR.

Fig. 1 shows the line-to-neutral voltage which must be sustained by load equipment and these values will remain realistic until reduced through component (other than OVR) design.

To the three basic sources of overvoltage enumerated by Mr. Stineman, we would like to add a fourth, namely, overvoltage caused by an open phase. The OVR should operate on categories 1 and 4.

Category 2 overvoltage is determined by characteristics of fault clearing components for a particular electric system. It is required that the OVR does not operate on category 2 faults unless fault isolation components fail to operate, at which time the OVR may isolate the system to prevent fire and other damage.

Mr. Owens questions the approach used in defining overvoltage. If we were not attempting to be realistic, we would be in agreement with him. As it is, however, we consider his suggestions to be unrealistic for the following reasons:

1. Mr. Owens believes that an overvoltage should be defined in relation to the value which utilization equipment can withstand without failure or malfunction. We agree and expressed this attitude in the paper and in reference 2. On many occasions, however, the electric power requirements of vital aircraft equipment is not known until well after the electric system has been designed, procured and installed. Reference 1 was issued, therefore, to inform utilization equipment manufacturers as to the characteristics of standard aircraft electric power.

2. We used the a-c generators, regulators, and other components available at the time when our overvoltage study was started. At least four of the seven a-c generators used were subsequently produced in significant quantity, and these generators define the maximum overvoltage values presented. Although there has been continuous improvement in a-c generators and regulators, these improvements, to our knowledge, have not yet significantly changed the parameters related to overvoltage. Again, production versions of new generators usually become available much too late to permit electric system redesign to account for minor variations in component characteristics.

3. The authors anticipate that electric system characteristics will be continuously improved, but do not agree that these characteristics should remain undefined until the ultimate a-c electric system is realized.

Calculations on Voltage Unbalance for 3-Phase Synchronous Systems

B. J. WILSON
ASSOCIATE MEMBER AIEE

W. K. GARDNER
NONMEMBER AIEE

THE steady-state behavior in regulated systems which become unbalanced has been treated in an earlier paper.[1] The treatment there was concerned principally with fault-type unbalances, though general unbalance was considered and amplified with an example. A need was felt for an extension in the direction of further quantitative evaluation of the degree of unbalance in voltage experienced as a function of load impedance unbalance in typical synchronous systems. The fact that in smaller systems the infinite bus concept no longer is tenable also calls for a simultaneous consideration of generator and load in affecting the final net bus line voltages. The work described here is a step in that direction, taking up general unbalance in wye-wye neutral-connected systems with illustrations for typical load conditions.

The limitations to the generality of analysis are those usually employed in symmetrical component analysis, namely, existence of linearity and exclusion from consideration of space and phasor harmonics other than the fundamental.

Background

Since the inception of the technique of symmetrical component linear transformation, much discussion has taken place assessing its relative worth as a tool in network analysis. Since the mathematical nature of symmetrical components is rather straightforward, these arguments have served at least to help orient into the over-all picture of polyphase circuitry those contributions provoking such discussions. This orientation aspect, so desirable, but so often lacking in technical periodical literature, should be emphasized.

In accord with this principle, the relation of the treatment presented here to earlier work on symmetrical components should be emphasized. The underlying features in obtaining these results can be described in terms of their contribution to the topological behavior and the speed and ease of solution to a high degree of accuracy for such unbalanced systems. The labor connected with programming an unbalance problem for effects of parametric variation has been responsible to a large extent for the scarcity of solutions of this type. Although the effects of the influence on unbalance of impedance variation are known to a limited extent in cases of common fault expressions, there does not seem to be a great amount of literature dealing with such solutions. As an example, a series of curves is presented by Wagner and Evans[2] illustrating the effects of symmetrical component impedance on fault current and voltages for a range of negative reactance restricted to typical values. However, the treatment is not extended to include more general cases of unbalance, where calculations generally are much more difficult than for the fault-type unbalances. Since Fortesque's[3] time, many works have appeared dealing with performance of machines under unbalanced conditions. Much work needs to be done to extend existing analyses and attempt to accomplish some sort of topological presentation of polyphase system performance under conditions of unbalance. In this way system designs can be fabricated from a more enlightened viewpoint on the limitations present due to unbalance and the minimum precautions to observe for satisfactory system operation.

For the circuits considered here, the manner in which the mathematical model is deduced from the given network configuration is greatly influenced by the process of solution to be used.[4,5] It is relatively easy to set up the circuit equations for the majority of unbalance problems. On occasion, some difficulty is experience in performing those transformations desired to facilitate final solutions by some specific process.[6,7] The elimination process in a system of equations, or what is tantamount thereto in the form of determinental solutions, usually constitutes the most time-consuming single step in the solution. In this age of equation solvers and computers, sidestepping this most difficult stage is relatively easy.

Hence the problem as described in the Appendix was formulated with a view to circumventing calculations with the labor and rounding error inherent in the solution of such systems of equations by longhand methods and under the added stipulation imposed by a considerable amount of parameter programming. An additional objective was to provide a guide in setting up for evaluation cases other than those exemplified in this report.

It was found advantageous to program this problem on an analogue computer of the mathematical model type, employing an iterative process for solution. This was decided in spite of the relatively low order of the systems of equations experienced.[8] Carrying out the algebraic form of problem solutions in matrix terminology resulted in a most convenient final form for subsequent numerical solutions via the computer. The relative ease with which the general solution degenerates to the particular cases can be observed from those examples cited in this paper.

Nomenclature

E = positive-sequence excitation voltage
a = algebraic operator $e^{j}\dfrac{2\pi}{3}$
I_0, I_1, I_2 = zero-, positive-, and negative-sequence currents
V_0, V_1, V_2 = zero-, positive-, and negative-sequence voltages
Z_0, Z_1, Z_2 = zero-, positive-, and negative-sequence generator impedance
Z_0', Z_1', Z_2' = zero-, positive-, and negative-sequence equivalent symmetrical component load self-impedances
Z_{aa}, Z_{bb}, Z_{cc} = phase self-impedances
Z_a, Z_b, Z_c = phase self-impedances
Z_{ab}, Z_{bc}, etc. = phase mutual impedances
Z_m = reference value of mutual impedance
K = constant $0<K\leq 1$
I_a, I_b, I_c = phase currents
V_a, V_b, V_c = phase voltages
I_{sc} = matrix representation of symmetrical component currents
V_{sc} = matrix representation of symmetrical component voltages
Z_{sc} = matrix representation of symmetrical generator impedance
Z = matrix representation of actual load impedances
Z' = matrix representation of transformed uncoupled load impedances
Z_c' = matrix representation of transformed coupled load impedances
Z_{scT} = sum of Z_{sc} and Z'
Z_{scT-c} = sum of Z_{sc} and Z_c'
$\| \ \|^{-1}$ = inverse of matrix $\| \ \|$
R = real part of impedance Re_z
X = imaginary part of impedance Im_z
ϕ = symbol for phase
θ = phase angle between phases

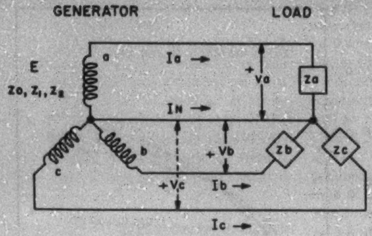
Calculation of Effect on Unbalance of Impedance Variations for Representative Loads

The wye-wye neutral-connected circuit is chosen here for the initial illustration

---

Paper 54-359, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted June 16, 1954; made available for printing August 6, 1954.

B. J. WILSON and W. K. GARDNER are with the Naval Research Laboratory, Washington, D. C.



GENERATOR CONSTANTS

ZERO-SEQUENCE IMPEDANCE, $Z_0=(0.071+j0.080)$ PER UNIT
POSITIVE-SEQUENCE IMPEDANCE, $Z_1=(0.052+j1.45)$ PER UNIT
NEGATIVE-SEQUENCE IMPEDANCE, $Z_2=(0.080+j0.213)$ PER UNIT

GENERATOR TEST SETTING

FREQUENCY = 400 CYCLES
EXCITATION VOLTAGE = POSITIVE-SEQUENCE VOLTAGE,
$E_1=E=(1+j0)$ PER UNIT

PER UNIT VOLTAGE AND CURRENT BASES ARE RATED VALUES FOR THE MACHINE.

Fig. 1. Wye-wye neutral-connected circuit

because of its widespread application. The specific examples of circuitry, i.e., resistive, inductive, coupled, etc., that fall under the wye-wye category are intended to be most representative of practical application in terms of power factor and expected degree of impedance unbalance. However, it is to be emphasized that the techniques employed in their solution generally are applicable to other circuit arrangements.

To facilitate graphical presentation of the unbalance behavior of the circuits, a sample generator was chosen in the form of a 400-cycle aircraft alternator with specific constants. The number of places that a quantity to be varied appears in a given closed form of solution makes it unfeasible to assign as a graphical variable some general parameter that includes arbitrary machine constants. The per-unit values that do appear in the calculations are sufficiently representative of practical cases to provide a close connection in these results with those expected in other systems. Furthermore, the characteristics of the analogue computer employed were such as to necessitate the use of numerical constants.

The method of obtaining equations which yield a solution to the wye-wye generator-load circuit follows that prescribed by symmetrical component analysis. It consists of writing Kirchhoff equations in terms of the symmetrical component quantities. It is not necessary, nor is it desirable, to use equivalent circuits in the usual symmetrical component fashion in the process of solution. Instead, the actual circuit under consideration is dealt with directly as it exists. This is not to say that a complicated load circuit would not be reduced to its true effective impedance, but rather that a solution can be obtained without resorting to artificial equivalent circuits. The properly written Kirchhoff equations form a set of simultaneous linear algebraic equations to be solved for voltages or currents, depending on the form, while the known quantities are the sequence impedances and generated voltage. In general, the choice of final voltage or current forms depends on the information desired, but may be dictated by the method used to solve the simultaneous equations. The analogue computer used in this investigation re-

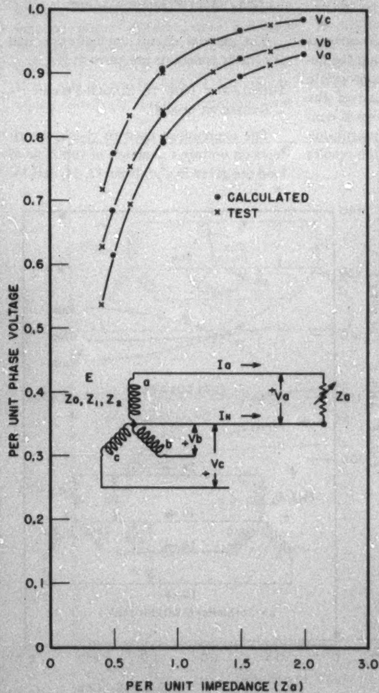


Fig. 2 (left). Effect of resistance variation on phase voltages for single-phase resistive load

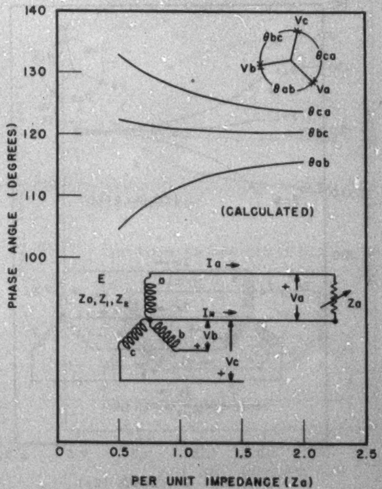


Fig. 3 (right). Effect of resistance variation on phase angles for single-phase resistive load

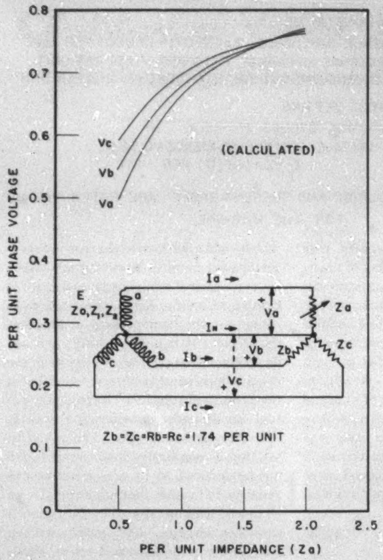


Fig. 4 (left). Effect of resistance variation on phase voltages for 3-phase resistive load

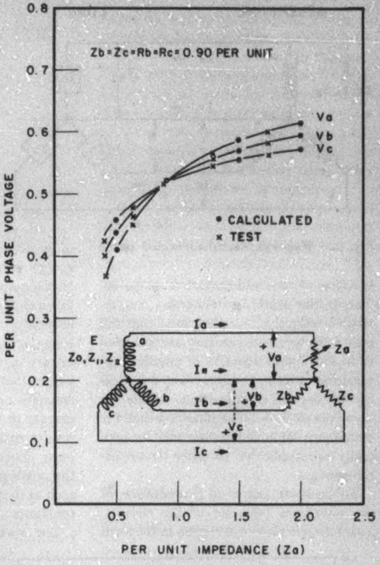


Fig. 6 (right). Effect of resistance variation on phase voltages for 3-phase resistive load

quired a form of equations satisfied only by the type where the symmetrical component currents appeared as unknowns. This fact is attributed to the properties of the particular circuits under investigation as they manifest themselves in an analagous mathematical manner. Thus, most of the sets of equations for the cases considered here were solved for currents which were used further to obtain the desired voltages by means of appropriate Kirchhoff relations. An outline of the development of these expressions is contained in the Appendix. Both magnitude and angle are extracted from the results for individual presentation.

The general circuit configuration and generator constants are given in Fig. 1.

UNBALANCE DUE TO SINGLE-PHASE RESISTIVE LOADS

The expressions used to check the effects on voltage unbalance of this type of load are given in equations 13, 14, and 15.

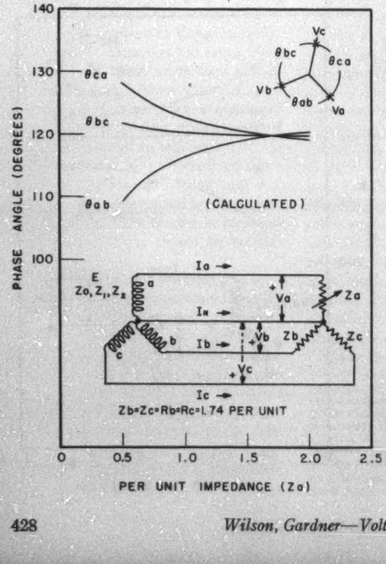


Fig. 5 (left). Effect of resistance variation on phase angles for 3-phase resistive load

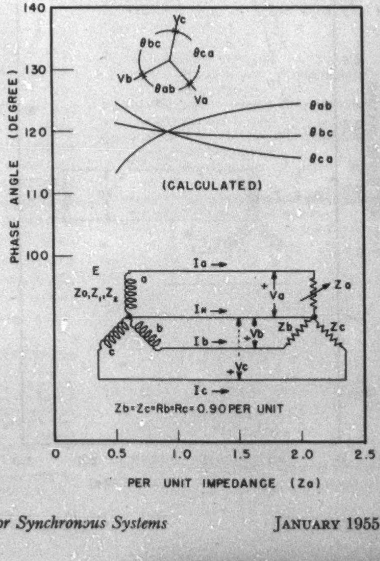


Fig. 7 (right). Effect of resistance variation on phase angles for 3-phase resistive load

Their form is dictated by the convenience to the solution it offers.

To approximate typically expected load conditions, Z_a (entirely resistive in this case) was varied from 0.5 to 2.0 per unit, while phases *b* and *c* remained open-circuited. The magnitude and angle of the resulting phase voltages are plotted correspondingly in Figs. 2 and 3. Results show close agreement between calculated and test voltages.

Unbalance Due to 3-Phase Resistive and Reactive Uncoupled Loads

Equations 11 provide the working expressions in evaluation of the unbalance effects of general 3-phase uncoupled loads. For the purely resistive load, and for the practical impedance variations of from 0.5 to 2.0 per unit, equations 11 remain diagonal. This fact satisfies the criterion for the iterative-type solution employed on equations 11. The form presented by equations 16 likewise allows solution by the same methods.

Calculations were made for a range of resistance of phase *a* from 0.5 to 2.0, with *b* and *c* phases remaining fixed at 0.40, 0.90, and 1.74 per unit while the *a*-phase impedance was varied. Final solutions for the phase voltages were achieved through transformation 4 and Ohm's law equation 5. The results of these calculations are shown in Figs. 4 through 9, which illustrate both magnitude and angle of phase voltages. The calculations were checked to close agreement by test for 0.90 per-unit impedance values in the phases whose resistance remain fixed.

For resistance-inductance impedances, equations 11 and the subsequent expansion to real form given by equations 16 yielded the solutions illustrated in Figs. 10 and 11. In this instance, the *a*-phase impedance was varied from 0.5 to 2.0 per unit at the constant 0.75 power factor. The *b*- and *c*-phase impedances remained at 1.5 per unit, with the same power factor. Fig. 10 illustrates the close agreement in test and calculated results.

Unbalance Due to Coupled 3-Phase Static Loads

For coupled circuits, the form of equations represented by 20 was employed in obtaining solutions. Combined with expression 18, the expanded form of these equations is written as shown in 21. Following the application of transformation 4 the final phase voltages are obtained through equations 23.

Two types of parameter variations were performed on the coupled form of 3-phase circuit. These were chosen largely on the basis of experimental convenience. In one type, the resistive component of the self-impedance of the *a*-phase was varied from 1.5 to 4.0 per unit. Calculations were performed also for an open-circuit condition on the *a* phase. Phases *b* and *c* were maintained at 1.5 per unit self-impedance with 0.75 power factor. Mutual coupling was equal among all phases and was chosen equal to one-half the reactive component of self-impedance. Equation 22 was employed to make the calculations. The results of these calculations for voltage magnitude are shown in Fig. 12 along with the confirmation by test of circuit performance. Calculations for the phase angles are given in Fig. 13.

In the other type of parametric variation, the amount of inductive coupling between one phase, the *a*-phase for reference, and the other two phases was changed incrementally. The subsequent calculations for this case provided one illustration of unbalance due to unequal coupling. In detail, the coupling of *a* phase to the other two was reduced from an equicoupled condition, designated by Z_M, by a factor of K, where $0<K\leq 1$. This was accomplished by removing $(1-K)$ turns from the coupled portion of the phase winding. The consequent coupling of the *a* phase with the others can be expressed as KZ_M per unit. The resultant self-reactance of *a* phase can be expressed with reasonable approximation as K^2X_{aa} in the event all turns of the *a* phase self-inductance were originally coupled to the other phases. For the conditions prescribed, the load impedance matrix becomes that shown in equation 24.

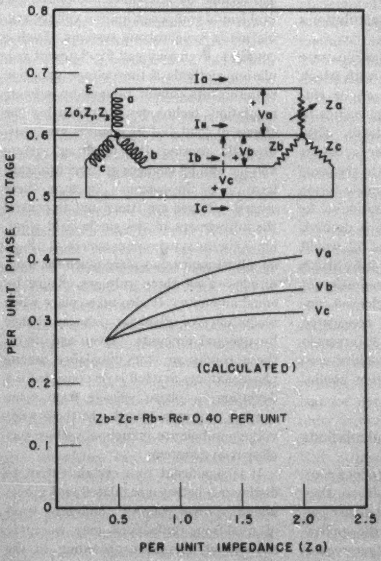
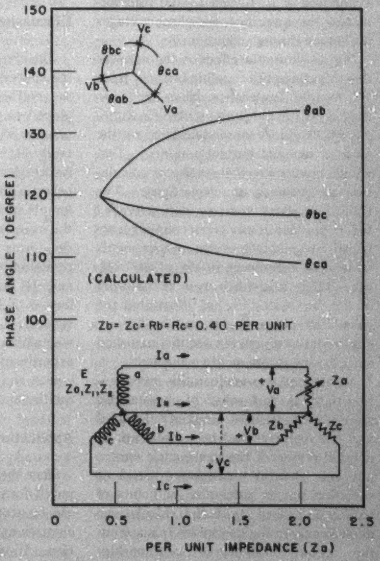


Fig. 8 (left). Effect of resistance variation on phase voltages for 3-phase resistive load

Fig. 9 (right). Effect of resistance variation on phase angles for 3-phase resistive load



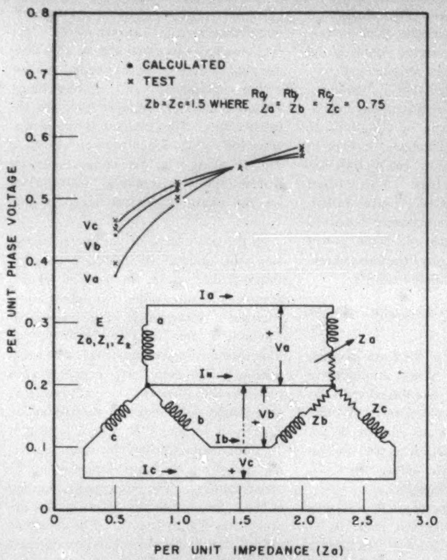


Fig. 10. Effect of impedance variation on phase voltages for 3-phase constant power-factor load

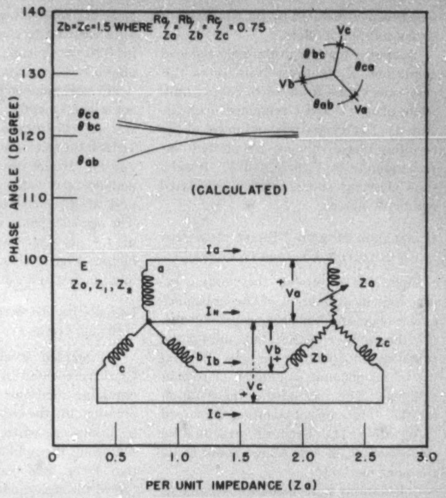


Fig. 11. Effect of impedance variation on phase angles for 3-phase constant power-factor load

Upon performing transformation 8 on equation 24 and inserting the results in equation 20, equation 25 is formed. Upon evaluation of I_0, I_1, and I_2, and transformation by equation 4, the phase voltages are solved through equation 23.

Fig. 14 shows the effect on the voltages due to varying the coupling factor K from 1/3 to 1.0. Impedance values are $Z_{aa}=Z_{bb}=Z_{cc}=1.5$ per unit, with $(Re\ Z_{aa})/Z_{aa}=0.75$ prior to modification to the state of unequal mutual coupling. Test values confirming the results of calculation are shown in the same figure. The three test phase voltages were averaged before plotting at any given coupling level because of inability of the measurements to resolve differences in the phase voltages, differences which were in the order of 0.2 per cent. Fig. 15 illustrates the calculated phase angles and in this instance only two curves are distinguished since θ_{ab} and θ_{bc} are nearly coincident.

No attempt was made in this investigation to obtain test values of phase angles since once agreement between calculated and test voltages had been obtained an over-all review of the phase-angle curves not only showed a consistent trend of variation but in particular, at points of balanced 3-phase load the calculations differed only in the hundredths place from the known 120-degree relationship. These factors in combination were considered a good indication of the reliability of the phase-angle calculations.

Discussion of Unbalance Calculations

The chosen examples serve to illustrate the success and relative ease with which the methods of calculation can be employed to yield phase voltage values in systems with unsymmetrical loads. Certainly it is not true that the equation forms selected for application to the computer in this instance are requisite forms for all such analogue computations. In the event digital computation is desired, the forms given in equations 12 would probably be preferred. Figs. 2, 6, 10, 12, and 14 demonstrate that accuracies of from 1 to 2 per cent can be achieved consistently. Such results are considered well within the limits imposed by errors in experimentally determined machine constants and neglect of saturation nonlinearities and harmonic effects.

Application of Unbalance Calculations

Once the phase voltages have been computed for an unbalanced condition, they can be used in completing unbalance calculation appropriate to a specific application. In instances where voltage regulation is of concern and perhaps detrimental to the operation of an electric system, it is possible to show how unbalance may further aggravate such effects. Incremental load changes in polyphase systems that do not produce equal increments of voltage change on all phases compound the problem of compensating for voltage regulation in synchronous systems. Such a condition even may call for different regulation methods if the voltage variation requirements cannot be met by existing regulation techniques. Ordinarily, the effects of load and its changes are measured by sensing the resulting system voltage and its changes at some arbitrary location in the system. In turn, these sensed voltages are translated to control the adjustment of the single field winding (s) in the synchronous source(s). Such an adjustment takes place when the magnitudes of all three voltages change by equal amounts. It also takes place when magnitudes of all phase voltages change by unequal amounts. Over and above these conditions, the unbalance among phase voltages in itself is responsible for a deviation in phase voltage from some average or nominal level. In these ways voltage unbalance influences voltage regulation on a system.

It is significant to a certain extent to distinguish between regulated and unregulated voltage sources. Different ways of qualifying unbalance may be more or less advantageous according to the

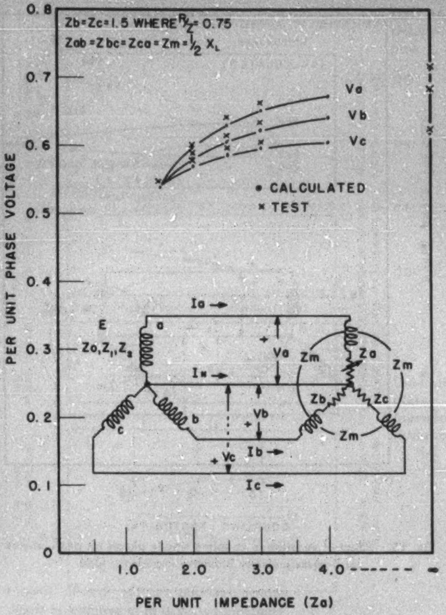


Fig. 12. Effect of resistance variation on phase voltages for 3-phase inductive impedance load with equimutual coupling

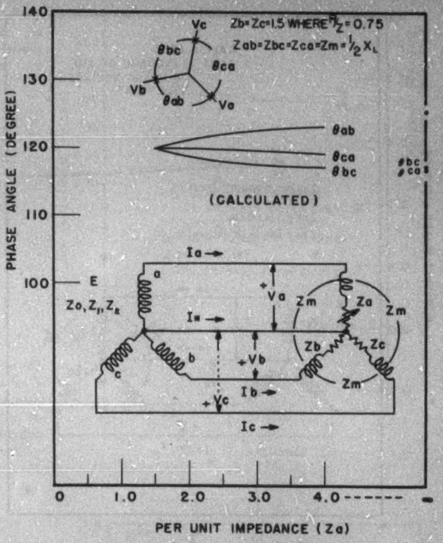


Fig. 13. Effect of resistance variation on phase angles for 3-phase inductive impedance load with equimutual coupling

manner in which equipment is affected by unbalance. For example, some rotating equipment is far more sensitive to unbalance than to voltage level, i.e., within reasonable bounds.[9] One definition for unbalance which is independent of level is that which describes a deviation in phase voltage from average in per cent of the average of the magnitudes of the unequal voltages. On the other hand, single-phase equipment operating on 3-phase systems obviously is little concerned with voltages other than its own. Here level is important. Since both types of equipment may operate on the same system, cognizance should be given both types of unbalance definitions.

The per-cent unbalance can be expressed as a ratio of the maximum deviation from the average of any phase voltage to the average. As a rough rule-of-thumb, such a definition has proved adequate, particularly for some military applications.[10] Employing this definition, per-cent unbalance can be extracted from Figs. 2 through 14. The results are illustrated in Figs. 16, 17, and 18. Besides prescribing impedance properties for a certain degree of unbalance, these curves permit evaluation of suitable application of equipment with known unbalance requirements or with known tolerable voltage variation requirements. Two specific illustrations of unbalance can be considered. When relatively small induction motors are applied to the systems for which calculations have been made, dissymmetries of 1 to 2 per cent would not be detrimental to their performance. However, unbalance of 5 to 6 per cent may reduce motor load limits by 25 to 30 per cent.[9] From Fig. 16(B) it is observed that at 1.2 per-unit resistive impedance on *a* phase and 1.74 per-unit resistive impedance on *b* and *c* phases, 2-per-cent unbalance occurs. At *a*-phase impedance less than 1.2 per unit, an unsatisfactory degree of unbalance exists for proper motor operation. As another specific example, military specifications[11] require of synchronous generators an unbalance not to exceed 4.0 per cent for unity power factor single-phase loads of 2/3 per unit kilovolt-ampere. In addition, displacement of adjacent phases of the fundamental should be within 115 to 125 degrees.[10] In Fig. 16(A) an impedance value of approximately 1.38 per unit corresponds to 2/3 per unit kilovolt-ampere. A value of slightly greater than 4-per-cent unbalance is observed at this impedance level. Phase displacements of 114, 120.5, and 125.5 are observed in angles between phases *a* and *b*, *b* and *c*, and *a* and *c* respectively.

In systems where voltage is regulated, generally it is more desirable to retain representation of the absolute magnitude of all phase voltages. Certainly no generality is sacrificed in this manner. In addition, a much clearer picture is gained of the actual spread among phases of the voltage, a fact which gives rise to a type of secondary regulation phenomenon. Hence proper consideration is gained of the effects of unbalance on both single- and 3-phase loads.

To present representative examples of the unbalance voltage spread in typical systems, the curves of Figs. 2 through 14 can be altered and replotted. Common regulating schemes attempt to hold constant some selected voltage, usually that of one phase or the approximate average of all the phases where allowances are taken for small amounts of voltage unbalance. The latter scheme will be chosen for illustration. By referring the average voltage at any abscissa value to a 1.0 per-unit level for the voltage curves in Figs. 2 through 14, the effect of having regulated to the average is achieved. The results of these transformations are illustrated in Figs. 16, 17, and 18.

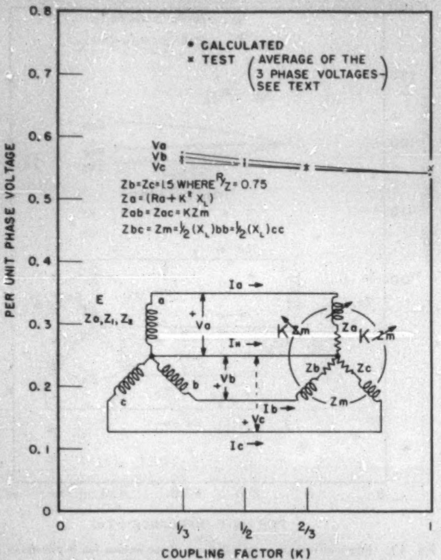


Fig. 14. Effect of variation in coupling among phases on phase voltages for 3-phase mutually coupled inductive impedance load

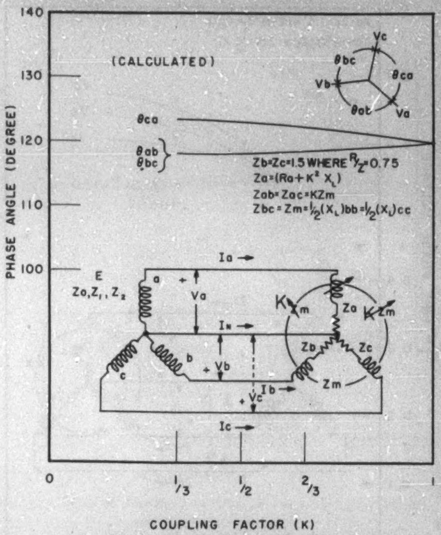


Fig. 15. Effect of variation in coupling among phases on phase angles for 3-phase mutually inductive impedance load

The regulated form of unbalance curves shown in Figs. 16, 17, and 18 emphasized more truly the secondary or quasiregulation effects that unbalance imposes on the system. The steady-state voltage of a system ordinarily has its regulator adjusted to maintain system voltages at certain levels under balanced conditions. Subsequent balanced voltage regulation usually is restricted to a specified limit. For purposes of illustration this limit will be assumed as ±2.5 per cent. Imposition of unbalance creates a departure in phase voltages from the nominal setting, which fact constitutes a quasiregulation effect. Further regulation occurs in the conventional manner upon variation of the load impedances. If the single-phase unbalance of Fig. 16(A) is considered, and if the regulator is assumed to maintain a constant average value, the mere addition of a 3/2 per-unit resistive impedance to *a* phase, following a no-load regulator adjustment to per-unit voltage, creates a quasiregulation of about 4 per cent on the *a* and *c* phases. A 1.0 per-unit impedance would create on these phases regulation of 5 and 6 per cent respectively. Although these particular results are not alarming in their specific contributions to regulation, they do exceed the assumed limits and are superimposed on all the others contributing to over-all system regulation, and as such may be important. The significance of this type of interpretation probably is less important to unit 3-phase loads than to individual single-phase loads. Nevertheless, it serves to illustrate an additional consideration that might advantageously be given in assessing over-all voltage regulation effects and provides a means for quantitatively evaluating them.

It is emphasized that calculations for unbalance were made on the basis of a single equivalent load. In reality the load can be either that of a single equipment or group of components possessing a unique effective value. By recognizing this qualification, it is possible to prevent certain misconceptions regarding the application of the results. In applying loads to systems for which unbalance calculations have been made, account must be taken of the system impedance changes introduced by the proposed addition of load. In other words, simultaneous solution of all circuit components in a Kirchhoff sense must be achieved before the degree of unbalance is known. This unbalance is then evaluated for comparison to what is required by the components themselves for suitable operation. In some instances strict adherence to this rule can be put aside. Voltage unbalance will be affected to a negligible extent by a load whose reaction on the system toward changing net system impedance is small. Hence, the existing degree of unbalance can be compared to that required by these "small" loads without further complication. Emphasis should also be given the role of the negative-sequence impedance Z_2 in influencing the degree of voltage unbalance at a given load impedance unbalance.

In many electric systems, the unbalance problem is resolved satisfactorily by proper distribution of loads or by the use of phase balancing devices. Resolution of unbalance in other systems without these auxiliarly corrective measures proves far more difficult. For example, in systems with relatively large intermittent loads, with equipment requiring an absolute minimum of unbalance, and especially with the necessity for achieving minimum over-all system weight, the usual methods prove inadequate.

It is apparent from the expressions developed and used in this investigation that if unbalanced loads are present the resulting unwanted voltage unbalance can be compensated to a great extent by modification of the generator sequence impedances, especially the negative-sequence impedance. Appreciation of the interrelated restrictions on these impedances as affected by considerations of practical machine design prevent over optimism in attempting such achieve-

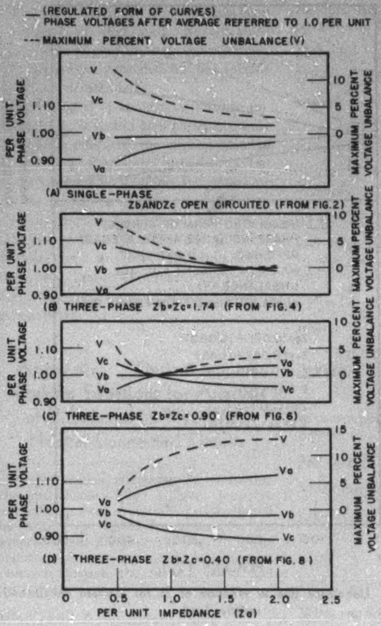
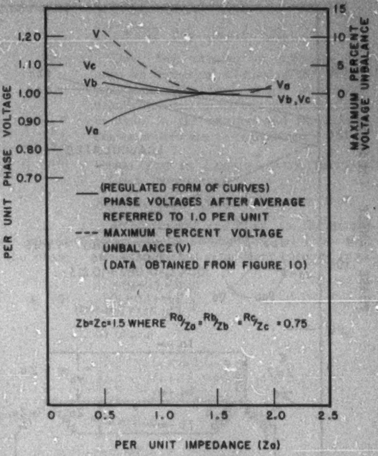


Fig. 16 (left). Unbalance for pure resistive load

Fig. 17 (right). Unbalance for 3-phase constant power-factor resistive and inductive load



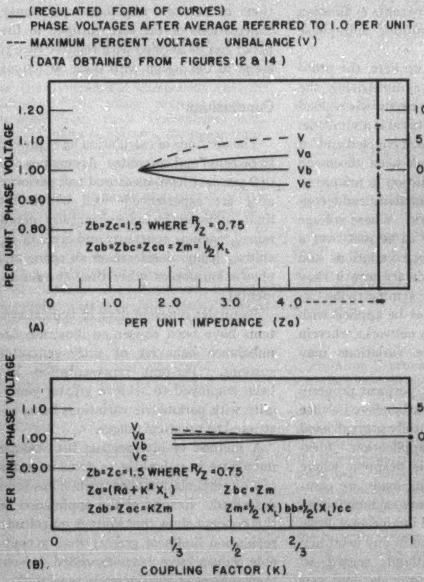


Fig. 18. Unbalance for 3-phase coupled load

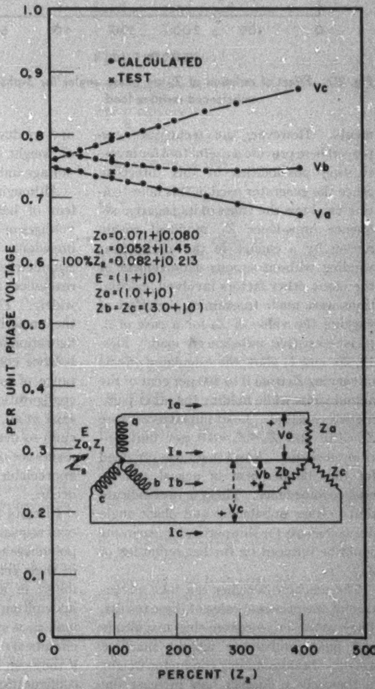


Fig. 19. Effect of variation of Z_2 on phase voltages for 3-phase unbalanced resistive load

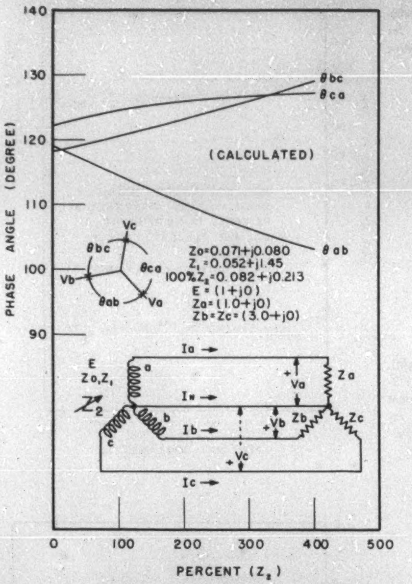


Fig. 20. Effect of variation of Z_2 on phase angles for 3-phase unbalanced resistive load

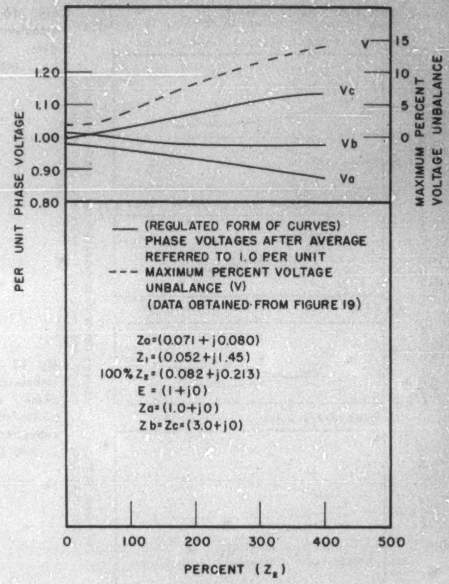


Fig. 21. Unbalance due to variation of Z_2 for 3-phase unbalanced resistive load

ments. However, the techniques employed here provide a useful tool for investigating possibilities in this direction. Since the generator used in this investigation can have the value of its negative-sequence impedance Z_2, modified within reason by a change in the amortisseur winding without serious derangement of the many other factors involved, calculations were made to examine the effect of varying the value of Z_2 for a case of 3-phase resistive unbalanced load. Figs. 19, 20, and 21 show the calculated effects of varying Z_2 from 0 to 400 per cent of the actual value while holding the other parameters constant. Load impedances were $Z_a=1.0$, and $Z_b=Z_c=3.0$ per unit, entirely resistive. A test point was obtained for the 100 per cent of actual zero-sequence impedance. These figures indicate that voltage unbalances and phase-angle displacements for this particular generator could be reduced by further reduction of Z_2.

The machine designer can take advantage of one principal value of these results. They assist in compromising magnitude and phase unbalance against machine weight. Besides adding conductor weight in themselves, damper bars increase the effective air gap, thus increasing excitation requirements. The link between Z_2 and conductor weight permits evaluation of weight versus magnitude and phase voltage unbalance.

Although not taken up here, the problem of balancing or symmetrizing the voltages of a polyphase circuit whose load impedance is unsymmetrical is a vital one. Symmetrizing by means of coupled and/or reactance networks has been discussed widely. The usual solution is unique to the prescribed load unbalance and configuration in the network. Where voltage balance conditions are to be met over a range of load impedance variations and configurations, solutions are rare if they exist at all. Obviously symmetrizing circuits so designed cannot be applied with assurance of success to networks wherein appreciable impedance variations may occur.

Circuits that do not perform properly over requisite parametric bands to balance polyphase voltages certainly are restricted in their generality of application. Flexibility of application is desirable where intermittent unbalancing loads are common in a system. Hence, symmetrizing circuits are most useful in a general sense if they can be connected in any arbitrary configuration, i.e. without regard to phase order at the terminals, and can perform satisfactorily over desired variations in network impedance unbalance. Calculations of the sort outlined in this paper can be applied with relative ease to assist in the design with these objectives.

Conclusions

The methods of calculation have proved to be rapid and accurate. Accuracies of 1 to 2 per cent from measured test performance are experienced—well within the limit imposed by experimentally determined machine constants and error in excluding from consideration all space and phasor harmonics other than the fundamental.

Examples repesentative of typical systems have been chosen to illustrate the unbalance behavior of single-generator systems. Per-unit representation has been employed to achieve greater generality with parametric variations being restricted to practical ranges.

A method of interpreting the voltage unbalance in polyphase systems in terms of a quasiregulation phenomenon has been described. Results of the application of this concept show that voltage magnitude regulation limits of greater than acceptable values have been exceeded in systems otherwise satisfactorily regulated.

The deviation in phase angle from $2\pi/n$

where n is the number of phases, has been calculated and illustrated. Results show that limits of greater than those acceptable have been exceeded for the specific applications intended.

The method of analysis of voltage unbalance employed in this paper not only provided a practical way of programming a variation of system impedances but it also serves as a guide for attacking similar problems in systems of different configurations. As general recognition increases of the capabilities of computing devices, the solution of such unbalanced systems no longer will present the tedious mathematical transformations formerly experienced.

Appendix. Derivation of Working Equations for Analysis of Unbalance in Wye-Wye Neutral-Connected Circuits

The expressions convenient for evaluating the performance of 3-phase wye-wye neutral-connected circuits can be formulated by writing the Kirchoff equations for the general circuit illustrated in Fig. 22. The equations for the symmetrical component voltages through the phase-a paths and neutral are

$$\begin{aligned} I_0Z_0+V_0&=0\\ -E+I_1Z_1+V_1&=0\\ I_2Z_2+V_2&=0 \end{aligned} \tag{1}$$

For simplicity neutral impedance will be assumed zero. Equations 1 contain a sufficient amount of information to solve for the symmetrical component currents indicated. The phase currents can be obtained from these symmetrical components solutions through the linear transformation relating them.

This fact can be illustrated most conveniently in the following manner by development through the use of matrix algebra. Initially the case of uncoupled loads will be treated. This is done to achieve simplicity in form without loss of generality. Inclusion of coupling, while requiring additional manipulative effort in development of final solution, does not materially alter this final form. This is shown in subsequent treatment of coupling effects. Proceeding from

$$\begin{Vmatrix}0\\E\\0\end{Vmatrix}=\begin{Vmatrix}Z_0&0&0\\0&Z_1&0\\0&0&Z_2\end{Vmatrix}\;\begin{Vmatrix}I_0\\I_1\\I_2\end{Vmatrix}+\begin{Vmatrix}V_0\\V_1\\V_2\end{Vmatrix}$$

or simply

$$\|E\|=\|Z_{sc}\|\;\|I_{sc}\|+\|V_{sc}\| \tag{2}$$

where the machine impedance is represented by

$$\|Z_{sc}\|=\begin{Vmatrix}Z_0&0&0\\0&Z_1&0\\0&0&Z_2\end{Vmatrix} \tag{3}$$

From the basic tenet of symmetrical components, the load phase voltages/currents are related to their symmetrical components by

$$\begin{aligned}\|V\|&=\|a\|\;\|V_{sc}\|\\ \|I\|&=\|a\|\;\|I_{sc}\|\end{aligned} \tag{4}$$

where

$$\|V\|=\begin{Vmatrix}V_a\\V_b\\V_c\end{Vmatrix};\;\|I\|=\begin{Vmatrix}I_a\\I_b\\I_c\end{Vmatrix};\;\|a\|=\begin{Vmatrix}1&1&1\\1&a^2&a\\1&a&a^2\end{Vmatrix}$$

Since

$$\|V_{sc}\|=\|a\|^{-1}\|V\|$$

and since the phase currents are related to the phase voltages through the phase impedances by

$$\|V\|=\|Z\|\;\|I\| \tag{5}$$

where

$$\|Z\|=\begin{Vmatrix}Z_a&0&0\\0&Z_b&0\\0&0&Z_c\end{Vmatrix} \tag{6}$$

in uncoupled 3-phase loads, equation 2 reduces to

$$\|E\|=\|Z_{sc}\|\;\|I_{sc}\|+\|a\|^{-1}\|Z\|\;\|a\|\;\|I_{sc}\|$$

which can be written

$$\|E\|=(\|Z_{sc}\|+\|Z'\|)\;\|I_{sc}\|=\|Z_{scT}\|\;\|I_{sc}\| \tag{7}$$

with

$$\|Z'\|=\|a\|^{-1}\|Z\|\;\|a\| \tag{8}$$

When expanded

$$\|Z'\|=\begin{Vmatrix}Z_0'&Z_2'&Z_1'\\Z_1'&Z_0'&Z_2'\\Z_2'&Z_1'&Z_0'\end{Vmatrix} \tag{9}$$

where

$$\begin{aligned}Z_0'&=\frac{1}{3}(Z_a+Z_b+Z_c)\\ Z_1'&=\frac{1}{3}(Z_a+aZ_b+a^2Z_c)\\ Z_2'&=\frac{1}{3}(Z_a+a^2Z_b+aZ_c)\end{aligned} \tag{10}$$

Substituting equations 3 and 9 into 7 there results

$$\begin{Vmatrix}0\\E\\0\end{Vmatrix}=\begin{Vmatrix}(Z_0+Z_0')&Z_2'&Z_1'\\Z_1'&(Z_1+Z_0')&Z_2'\\Z_2'&Z_1'&(Z_2+Z_0')\end{Vmatrix}\;\begin{Vmatrix}I_0\\I_1\\I_2\end{Vmatrix} \tag{11}$$

Equation 7 provides the solution for 1, and the phase currents can be obtained from 4. Rather than combine these steps and obtain a final explicit expression for the phase currents, it was preferred to leave the working equation in the form shown. By this expedient a diagonal form of equations was retained necessary for solution in an analogue computer employing the Gauss-Seidel iteration process. An alternative method of solving equation 7 for the currents involves the use of the inverse of the impedance matrix.

The expressions for phase currents and voltages, as obtained by inspection from equations 4, 5, and 7 are

$$\begin{aligned}\|I\|&=\|a\|\;\|Z_{scT}\|^{-1}\|E\|\\ \|V\|&=\|Z\|\;\|a\|\;\|Z_{scT}\|^{-1}\|E\|\end{aligned} \tag{12}$$

Thus, two avenues of approach are open for numerical solution of the voltage equations. The final choice of method usually depends on the equipment available to assist in the solution and the desired degree of accuracy to be achieved.

UNBALANCE IN CIRCUITS WITH SPECIFIC LOAD IMPEDANCES

Resistive Load—Single Phase

A convenient reduction to the single phase can be accomplished by expanding the symmetrical component currents from

$$\begin{aligned}\|I\|&=\|a\|\;\|I_{sc}\|\\ \|I_{sc}\|&=\|a\|^{-1}\|I\|\end{aligned}$$

to obtain

$$\begin{Vmatrix}I_0\\I_1\\I_2\end{Vmatrix}=\frac{1}{3}\begin{Vmatrix}I_a+I_b+I_c\\I_a+aI_b+a^2I_c\\I_a+a^2I_b+aI_c\end{Vmatrix}$$

Letting

$$I_b=I_c=0$$

equation 11 can be written

$$\begin{Vmatrix}0\\3E\\0\end{Vmatrix}=\begin{Vmatrix}Z_0+Z_0'&Z_2'&Z_1'\\Z_1'&Z_1+Z_0'&Z_2'\\Z_2'&Z_1'&Z_2+Z_0'\end{Vmatrix}\;\begin{Vmatrix}I_a\\I_a\\I_a\end{Vmatrix}$$

whose equations add to

$$3E=I_a[3(Z_0'+Z_1'+Z_2')+(Z_0+Z_1+Z_2)]$$

which reduces to

$$3E=I_a[3Z_a+(Z_0+Z_1+Z_2)]$$

As a result of this

$$V_a=\frac{3EZ_a}{3Z_a+(Z_0+Z_1+Z_2)} \tag{13}$$

By combining equations 2 and 4 and selecting the first and third equations of the result, there follows

$$0=V_a\left(\frac{Z_0}{Z_a}+1\right)+V_b+V_c$$

$$0=V_a\left(\frac{Z_2}{Z_a}+1\right)+a^2V_b+aV_c$$

from which

$$V_b=(a-a^2)\,\frac{V_a}{3Z_a}[a(Z_0+Z_a)-(Z_2+Z_a)] \tag{14}$$

Also from the same pair

$$V_c=-\frac{V_a}{Z_a}(Z_0+Z_a)-V_b \tag{15}$$

Resistive and Reactive Uncoupled Load—Three Phase

The reduction from equation 11 amounts to a mere substitution of impedance values. The result is a set of third-order algebraic equations, with complex variables and coefficients. If desired, a determinental solution can be effected immediately at this stage. However, it is desirable often to further alter the equations and form a 6 × 6 set all of whose elements are real. In such a case, equation 11 becomes

$$\begin{Vmatrix}0\\0\\E\\0\\0\\0\end{Vmatrix}=\begin{Vmatrix}(Z_0+Z_0')_r & -(Z_0+Z_0')_i & Z_{2_r}' & -Z_{2_i}' & Z_{1_r}' & -Z_{1_i}'\\(Z_0+Z_0')_i & (Z_0+Z_0')_r & Z_{2_i}' & Z_{2_r}' & Z_{1_i}' & Z_{1_r}'\\Z_{1_r}' & -Z_{1_i}' & (Z_1+Z_0')_r & -(Z_1+Z_0')_i & Z_{2_r}' & -Z_{2_i}'\\Z_{1_i}' & Z_{1_r}' & (Z_1+Z_0')_i & (Z_1+Z_0')_r & Z_{2_i}' & Z_{2_r}'\\Z_{2_r}' & -Z_{2_i}' & Z_{1_r}' & -Z_{1_i}' & (Z_2+Z_0')_r & -(Z_2+Z_0')_i\\Z_{2_i}' & Z_{2_r}' & Z_{1_i}' & Z_{1_r}' & (Z_2+Z_0')_i & (Z_2+Z_0')_r\end{Vmatrix}\begin{Vmatrix}I_{0_r}\\I_{0_i}\\I_{1_r}\\I_{1_i}\\I_{2_r}\\I_{2_i}\end{Vmatrix} \qquad (16)$$

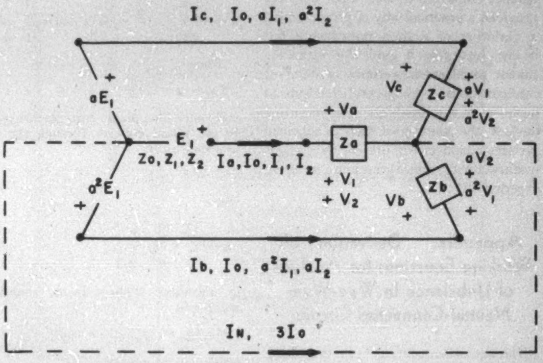


Fig. 22 (below). Direction conventions for the wye-wye neutral-connected circuit

where subscript r designates the real components and subscript i designates the imaginary components.

Final solutions for the phase voltages are achieved through transformation 4 and the Ohm's law equation 5.

Coupled 3-Phase Static Loads

The form presented by equation 7 is applicable also for instances of coupling within static loads. In such cases the Z' impedance factor assumes a more general role.

Proceeding from transformation 8 where

$$\|Z\|=\begin{Vmatrix}Z_{aa} & Z_{ab} & Z_{ac}\\Z_{ba} & Z_{bb} & Z_{bc}\\Z_{ca} & Z_{cb} & Z_{cc}\end{Vmatrix}=\begin{Vmatrix}Z_{aa} & Z_{ab} & Z_{ac}\\Z_{ab} & Z_{bb} & Z_{bc}\\Z_{ac} & Z_{bc} & Z_{cc}\end{Vmatrix} \qquad (17)$$

by assumption there results the following equation

$$\|Z_c'\|=\begin{Vmatrix}\left[Z_0'+\frac{2}{3}(Z_{ab}+Z_{bc}+Z_{ac})\right] & \left[Z_2'-\frac{1}{3}(aZ_{ab}+Z_{bc}+a^2Z_{ac})\right] & \left[Z_1'-\frac{1}{3}(a^2Z_{ab}+Z_{bc}+aZ_{ac})\right]\\\left[Z_1'-\frac{1}{3}(a^2Z_{ab}+Z_{bc}+aZ_{ac})\right] & \left[Z_0'-\frac{1}{3}(Z_{ab}+Z_{bc}+Z_{ac})\right] & \left[Z_2'+\frac{2}{3}(aZ_{ab}+Z_{bc}+a^2Z_{ac})\right]\\\left[Z_2'-\frac{1}{3}(aZ_{ab}+Z_{bc}+a^2Z_{ac})\right] & \left[Z_1'+\frac{2}{3}(a^2Z_{ab}+Z_{bc}+aZ_{ac})\right] & \left[Z_0'-\frac{1}{3}(Z_{ab}+Z_{bc}+Z_{ac})\right]\end{Vmatrix} \qquad (18)$$

All the first terms of equation 18 are of the same form as equations 10 in the similar transformation for the uncoupled case. These are

$$Z_0'=\frac{1}{3}\,[Z_{aa}+Z_{bb}+Z_{cc}]$$

$$Z_1'=\frac{1}{3}\,[Z_{aa}+aZ_{bb}+a^2Z_{cc}] \qquad (19)$$

$$Z_2'=\frac{1}{3}\,[Z_{aa}+a^2Z_{bb}+aZ_{cc}]$$

Hence, for a static coupled load

$$\|E\|=[\|Z_{sc}\|+\|Z_c'\|]\ \ \|I_{sc}\|=\|Z_{scT-c}\|\ \ \|I_{sc}\| \qquad (20)$$

Combined with expression 18 the expanded form of these equations is written

$$\begin{Vmatrix}0\\E\\0\end{Vmatrix}=\begin{Vmatrix}\left[(Z_0+Z_0')+\frac{2}{3}(Z_{ab}+Z_{bc}+Z_{ac})\right] & \left[Z_2'-\frac{1}{3}(aZ_{ab}+Z_{bc}+a^2Z_{ac})\right] & \left[Z_1'-\frac{1}{3}(a^2Z_{ab}+Z_{bc}+aZ_{ac})\right]\\\left[Z_1'-\frac{1}{3}(a^2Z_{ab}+Z_{bc}+aZ_{ac})\right] & \left[(Z_1+Z_0')+\frac{1}{3}(Z_{ab}+Z_{bc}+Z_{ac})\right] & \left[Z_2'+\frac{2}{3}(aZ_{ab}+Z_{bc}+a^2Z_{ac})\right]\\\left[Z_2'-\frac{1}{3}(aZ_{ab}+Z_{bc}+a^2Z_{ac})\right] & \left[Z_1'+\frac{2}{3}a^2Z_{ab}+Z_{bc}+aZ_{ac})\right] & \left[(Z_2+Z_0')-\frac{1}{3}(Z_{ab}+Z_{bc}+Z_{ac})\right]\end{Vmatrix}\begin{Vmatrix}I_0\\I_1\\I_2\end{Vmatrix} \qquad (21)$$

For equimutual coupling Z_m between phases, equation 21 reduces to

$$\begin{Vmatrix}0\\E\\0\end{Vmatrix}=\begin{Vmatrix}(Z_0+Z_0'+2Z_m) & Z_2' & Z_1'\\Z_1' & (Z_1+Z_0'-Z_m) & Z_2'\\Z_2' & Z_1' & (Z_2+Z_0'-Z_m)\end{Vmatrix}\begin{Vmatrix}I_0\\I_1\\I_2\end{Vmatrix} \qquad (22)$$

Again to achieve the final phase voltages an expression similar to 5 with $Z\|$ defined as in equation 17 is employed following application of transformation 4.

With mutual coupling present, this transformation becomes

$$\begin{Vmatrix}V_a\\V_b\\V_c\end{Vmatrix}=\begin{Vmatrix}Z_{aa} & Z_{ab} & Z_{ac}\\Z_{ab} & Z_{bb} & Z_{bc}\\Z_{ac} & Z_{bc} & Z_{cc}\end{Vmatrix}\begin{Vmatrix}I_a\\I_b\\I_c\end{Vmatrix} \qquad (23)$$

There is no great complication by the additional considerations of neutral impedance and neutral coupling, the effect being to add terms to those of expression 18. The most widely encountered effect of neutral impedance, i.e., self-impedance, in a neutral circuit is handled by the addition of thrice the neutral self-impedance to the term in the first row and column of 18.

To illustrate conveniently the effects on unbalance of coupling through a static circuit, certain conditions were prescribed on equation 20 which could be duplicated conveniently in test. By assuming in addition to

$Z_{t_1} = Z_{f_1}$

that in per unit

$Z_{ab} = Z_{ac} = KZ_m;\; Z_{bc} = Z_m$

an expression can be formed which is a model for the situation where the coupling of one phase is varied with respect to the other two. For a typical reference let

$Z_m = \frac{1}{2}\, Im Z_{bb}$ per unit

As a result of these assumptions

$$\|Z\| = \begin{Vmatrix} Z_{aa} & KZ_m & KZ_m \\ KZ_m & Z_{bb} & Z_m \\ KZ_m & Z_m & Z_{cc} \end{Vmatrix} \qquad (24)$$

where

$Z_{aa} = Re Z_{aa} + jK^2\, Im Z_{aa}$

which transforms to

$$\|Z_c'\| = \begin{Vmatrix} \left[Z_0' + \frac{2}{3}(2K+1)Z_m\right] & \left[Z_2' + \frac{1}{3}(K-1)Z_m\right] & \left[Z_1' + \frac{1}{3}(K-1)Z_m\right] \\ \left[Z_1' + \frac{1}{3}(K-1)Z_m\right] & \left[Z_0' - \frac{1}{3}(2K+1)Z_m\right] & \left[Z_2' + \frac{2}{3}(1-K)Z_m\right] \\ \left[Z_2' + \frac{1}{3}(K-1)Z_m\right] & \left[Z_1' + \frac{2}{3}(1-K)Z_m\right] & \left[Z_0' - \frac{1}{3}(2K+1)Z_m\right] \end{Vmatrix} \qquad (25)$$

Incorporated into equation 20 the expanded form of these equations is written

$$\begin{Vmatrix} 0 \\ E \\ 0 \end{Vmatrix} = \begin{Vmatrix} \left[(Z_0+Z_0') + \frac{2}{3}(2K+1)Z_m\right] & \left[Z_2' + \frac{1}{3}(K-1)Z_m\right] & \left[Z_1' + \frac{1}{3}(K-1)Z_m\right] \\ \left[Z_1' + \frac{1}{3}(K-1)Z_m\right] & \left[(Z_1+Z_0') - \frac{1}{3}(2K+1)Z_m\right] & \left[Z_2' + \frac{2}{3}(1K-)Z_m\right] \\ \left[Z_2' - \frac{1}{3}(K-1)Z_m\right] & \left[Z_1' + \frac{2}{3}(1-K)Z_m\right] & \left[(Z_2+Z_0') - \frac{1}{3}(2K+1)Z_m\right] \end{Vmatrix} \begin{Vmatrix} I_0 \\ I_1 \\ I_2 \end{Vmatrix} \qquad (26)$$

References

1. METHODS FOR PREDICTION OF STEADY-STATE PERFORMANCE FOR UNBALANCED REGULATED 3-PHASE GENERATORS, **B. J. Wilson.** *AIEE Transactions*, vol. 72, pt. II, 1953 (Jan. 1954 section), pp. 413–22.

2. SYMMETRICAL COMPONENTS (book), **C. F. Wagner, R. D. Evans.** McGraw-Hill Book Company, Inc., New York, N. Y., 1933, pp. 228–35.

3. METHOD OF SYMMETRICAL CO-ORDINATES APPLIED TO THE SOLUTION OF POLYPHASE NETWORKS, **C. L. Fortescue.** *AIEE Transactions*, vol. 37, pt. II, 1918, pp. 1027–1140.

4. TWO-PHASE CO-ORDINATES OF A THREE-PHASE CIRCUIT, **Edward W. Kimbark.** *AIEE Transactions*, vol. 58, 1939, pp. 894–910.

5. TRANSFORMATION THEORY OF GENERAL STATIC POLYPHASE NETWORKS, **Louis A. Pipes.** *AIEE Transactions (Electrical Engineering)*, vol. 59, Feb. 1940, pp. 123–28.

6. LINEAR TRANSFORMATIONS IN 3-PHASE CIRCUITS, **Louis A. Pipes.** *AIEE Transactions*, vol. 60, 1941, pp. 351–56, discussion 619–21.

7. RELATIONS AMONG TRANSFORMATIONS USED IN ELECTRICAL ENGINEERING PROBLEMS, **Charles Concordia.** *General Electric Review*, Schenectady, N. Y., vol. 41, 1938, pp. 323–25.

8. LINEAR COMPUTATIONS, **P. S. Dwyer.** John Wiley and Sons, Inc., New York, N. Y., 1951, pp. 332–34.

9. LOAD CAPACITY OF INDUCTION MOTORS CONNECTED TO UNBALANCED THREE-PHASE SUPPLY, **R. Taschek.** *Elektrotechnika*, Budapest, Hungary, vol. 46, Jan. 1953, pp. 18–23.

10. *Military Specification MIL-E-7894*, Sections 3.1.2.2, 3.2.1.2.4, 4.1.1, 4.1.3, April 28, 1952.

11. *Military Specification MIL-G-6099*, Sections 4.5.6, 4.5.7, 4.5.7.1, September 1, 1948.

12. SYMMETRICAL COMPONENT ANALYSIS OF UNSYMMETRICAL POLYPHASE SYSTEMS, **R. Neumann.** Isaac Pitman and Sons, Ltd., London, England, 1939.

◆

Discussion

Bert V. Hoard (Boeing Airplane Company, Seattle, Wash.): The information presented in this paper is a valuable contribution in helping to solve the various problems arising from unbalanced electric loads on airplanes. The alternator selected for the tests appears to have impedances approaching the average of typical 400-cycle airplane types. Therefore the results of the paper in curve form can be applied directly to obtain rapid approximate solutions for specific problems without resorting to actual calculations.

It is of interest to refer briefly to the normal method of solving this type of problem, assuming a digital or analogue computer is not available or the problem has not been programmed on such a machine. On page 22, Fig. 19-h of reference 1 of the discussion, a double line-to-ground fault is shown and the solution equations are given on page 21, equations 34 through 39. A solution of a double line-to-ground fault normally is considered not difficult. Now refer to page 22, Fig. 19-k, which involves unbalanced star loads with grounded neutral similar to those solved for Figs. 2 through 10 in the paper being discussed. Figs. 19-h and -k are very similar except for the introduction in Fig. 19-k of the load impedances Z_b and Z_a in series in the various networks, and the impedance branch $1/3(Z_a - Z_b)$ paralleling the three resultant networks. As a result of this similarity the unbalanced impedance case is not much more difficult to solve than the double line-to-ground fault case.

Where only one or two check points are required this method of solution may be faster, and it is easier to follow unless the engineer is well versed in determinants or matrix mathematics. It is realized that where a large number of unbalanced cases must be calculated, as in the paper, the determinant or matrix method as used in the paper will save time, especially if a digital or analogue computer of the mathematical-model type is avail able.

REFERENCE

1. ELECTRICAL TRANSMISSION AND DISTRIBUTION BOOK (book). Westinghouse Electric Corporation East Pittsburgh, Pa., 3rd ed., 1944.

B. J. Wilson and **W. K. Gardner**: Mr. Hoard's comments provide a very useful and cogent supplement to the paper. It is gratifying to learn that the results have some general applicability to the current series of aircraft alternators.

The observation made by Mr. Hoard to the effect that the utility of the method is greatest where more than just a few check point calculations are required generally is true for the common fault unbalances. There are also some types of unbalanced systems for which closed forms are simple enough to make longhand calculations practicable. However, such instances occur usually when one or more of the system variables lose their dependency or have an evident interrelationship. When all the variables remain, full blown and independent, the solution of the 6 × 6 system (or greater, depending on the number of phases) in longhand form is difficult by standards of the average engineer. On the other hand, calculation of the unbalance problem on any occasion for few or many cases is a simple matter once the machine calculating routine or technique is learned.

Current Distribution in Paired-Phase Bus Bars Under Unbalanced Load Conditions

J. B. CATALDO
ASSOCIATE MEMBER AIEE

N. SHACKMAN
ASSOCIATE MEMBER AIEE

THE operating characteristics of feeder systems for the distribution of large polyphase currents at low voltage are usually given on the basis of performance when conducting balanced steady-state currents. Voltage drops and temperature rise are measured, and impedance calculated, from simple tests involving the passage of balanced polyphase currents through one or more sections of the feeder duct having a wye tie at the end of the last section. The results are averaged and used to describe the rating of the system.

Although such a standard of test is probably necessary for the comparison of performance between devices or pieces of equipment, the results are not always completely applicable or adequate, their use depending a great deal on the extent to which they reflect performance under actual service conditions. With feeder systems, the balanced conditions under which they are tested are rarely or, at best, only momentarily encountered in actual service. Loads and power factors are continually changing so that an unbalanced, rather than a balanced, condition prevails. The efficient distribution of power under all variations of service then becomes one of the most important qualities of the feeder system.

One of the most important characteristics contributing to efficient performance in multibar feeder systems is the degree of uniformity with which the phase currents distribute themselves among the bus bars. Although maximum uniformity of distribution probably occurs under balanced load conditions, widely unequal distribution under other conditions as, for example, with large single-phase loads, can promote undue heating in some of the bars and an increase in stray magnetic fields. Such effects limit the efficient use of the conductor material and seriously decrease the extent to which the system may be used under overload conditions. A study has been made under both field and laboratory conditions to determine the manner in which currents distribute themselves in the bus bars of a multibar feeder system under various conditions of load unbalance.

Feeder System Characteristics

The design of feeder systems or busways is concerned generally with the geometrical arrangement or configuration of bus bars to permit optimum characteristics of voltage drop and temperature rise. The problem is one of arranging the bus bars so that the alternating currents will flow with as uniform a current density as possible throughout the conductor areas for the most efficient utilization of the conductor material. To accomplish this, most systems employ two or more bus bars per phase in close spacings to permit reduction of skin effect and reactance voltage drops. However, the choice of the number of bars per phase, their physical and electrical arrangement, and the closeness of spacings must be viewed from a practical standpoint, so that ease of mechanical installation and electrical connections is obtained.

Two of the most important considerations in feeder system design are voltage drop and current-carrying capacity. The need for maintaining voltage drop at a minimum is apparent and it can be said, generally, that most systems are successful in this respect. Comparing the differences in voltage drops of various types of configurations with their operating voltages readily demonstrates this fact.

The second characteristic, that of current-carrying capacity, is equally as important as voltage drop but has a wider variation with different bus bar arrangements. Capacity is determined by the size, shape, thermal characteristics, and geometric arrangement of the conductor material, and is based on some maximum permissible temperature rise. In considering this characteristic, uniform current densities and equal distribution of current among bus bars are desirable so that the resultant maximum temperature rise of any bar may be as close as possible to the average of all bus bars in the system. The obvious result is the efficient use of all of the conductor material.

Although a maximum permissible temperature rise is specified in rating the current-carrying capacity of a duct, most systems are standardized on a given conductor cross-sectional area per phase for a given current rating. As a result, wide variations can exist among various feeder configurations both in maximum temperature rise and in maximum to minimum rise among the bus bars. These variations are caused generally by the degree of inequality of current distribution in and among the bus bars and to the thermal dissipation characteristic of the system as a whole. Essentially, therefore, current distribution is of prime interest in considering the current-carrying capacity of a feeder system.

Characteristics of Feeder Duct Under Test

The feeder duct with which the series of current distribution tests were conducted employs two conductors per phase in an arrangement appropriately described as paired phase.[1] The six conductors of the duct are arranged in three separate and distinct pairs, each pair consisting of two wide bars in close face-to-face proximity. Each bar of the pair is connected to a different phase. The pairs are arranged in edgewise proximity as shown in the diagram of Fig. 1.

Characteristics of this system include extremely uniform current densities with a consequent maximum reduction in skin effect, minimum resistance and reactance voltage drops, and almost perfectly balanced line-to-line impedances. Low a-c energy losses are obtained which, combined with effective heat dissipation, produce low and uniform temperature rises in each bar.

A peculiar inherent characteristic of the paired phase configuration is that the currents in each closely spaced pair of bars are nearly equal in magnitude and opposite in phase. This effect is similar to that of magnitude and phase distribution of currents in a single-phase transmission system, even though each conductor of

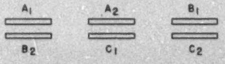


Fig. 1. Paired-phase bus bar arrangement

Paper 54-329, recommended by the AIEE Industrial Power Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing June 30, 1954.

J. B. CATALDO and N. SHACKMAN are with the Bulldog Electric Products Company, Detroit, Mich.

The authors wish to acknowledge the assistance, in the preparation of this paper, of Dr. Paul Hoover, head of the Electrical Engineering Department, Case Institute of Technology.

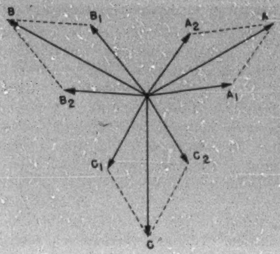


Fig. 2. Bar and phase current distribution for balanced loads

the pair is connected in parallel with another conductor in a different pair. As shown in Fig. 2, each of the three phase currents A, B, C, divides into two bar components (A_1 and A_2, B_1 and B_2, C_1 and C_2) which are nearly equal in magnitude and which differ in phase angle by nearly 60 degrees. Thus, the currents in each pair of bars (A_1 and B_2, A_2 and C_1, B_1 and C_2) are essentially equal in magnitude and opposite in phase.

The effects of this current distribution among the bars are uniform current densities, the practically complete neutralization of the magnetic fields of the two equal and opposite currents, and negligible eddy current losses in the magnetic enclosures and surroundings. It is also obvious that equal distribution of currents will produce more uniform temperature rises among bars.

The investigation and evaluation of these characteristic for balanced load conditions have been previously reported by Fisher and Frank.[1] The maintenance of this same characteristic under unbalanced load conditions is of equal importance and interest for optimum performance of the duct. If the physical arrangement of the system were such as to force the currents to distribute themselves in essentially equal and opposite relationships in bar pairs, a situation can exist which will tend to reduce the spread of current magnitudes among the bus bars under all load conditions.

Methods of Tests

The simultaneous determination of bus bar current magnitudes and phase relationships has required a departure from techniques employing standard current transformers because of close bus bar spacings. The arrangement of the bus bars in the paired-phase system makes it possible to maintain the designed close spacings of the bars at all joints, bends, and other fittings. Special bends or offsets would therefore have been required to employ standard current transformers in the tests which, it was felt, might introduce measurement errors. Convenient use was therefore made of a Rogowski coil,[2,3] which is essentially an air core solenoid of many turns wound around a nonconductor, the assembly being wrapped around the bar to be measured. Its thin, wide cross-sectional area makes possible its insertion between bus bars without disturbing bus bar spacings.

Six Rogowski coils were constructed and checked for phase and magnitude linearity. The coils were then calibrated and paired with six magnetic elements of a recording oscillograph to make possible the simultaneous measurement of the characteristics of all bus bars. Fig. 3 shows the Rogowski coils in position on the duct.

Voltage drop measurements were also taken with an oscillograph. Power loss was measured with a special low-voltage wattmeter, and temperatures by means of thermocouples.

Factory Tests and Results

Because greatest interest is naturally associated with actual field conditions, qualitative tests were first conducted on a portion of a feeder run installed in an industrial plant. The run fed a combination of single-phase and 3-phase infrared lamp loads, and a single-phase welder. Measurements were made on a 30-foot section which included a conventional elbow and began at a distribution transformer. The 3-phase duct was rated at 1,350 amperes, 600 volts, and consisted of two 1/4- by $2^{7}/_{16}$ inch copper bars per phase. The housings were ventilated and each bus bar was wrapped with varnished cambric tape. Interphase ties were made at both the line and load ends of the duct. These consisted of tie bars connecting the phase bars together, such as A_1 and A_2, B_1 and B_2, and C_1 and C_2. Fig. 4 shows a view of the duct under discussion.

For visual analysis purposes, adjacent oscillograph elements were connected to paired bars in accordance with the arrangement of Fig. 1. Thus, phase relationships can be readily observed between paired bars A_1 and B_2, A_2 and C_1, and B_1 and C_2.

The oscillographic records of the tests are as follows.

1. Fig. 5 shows the results of a 3-phase balanced infrared lamp load of approximately 400 amperes.
2. Fig. 6 illustrates the division of current among all bars for a single-phase infrared lamp bank, 400 amperes, across phase A-C.
3. Fig. 7 shows the change in bar currents when a welder load (40 kva) is added to an existing 400-ampere 3-phase infrared lamp load. The load was added to phase A-C.
4. Fig. 8 illustrates the effect on all bars by the addition of the welder load to an existing 400-ampere single-phase infrared lamp load, all on phase A-C.
5. Fig. 9 clearly shows the effect of transient conditions on the equal and opposite current characteristics. The load is the same as that of Fig. 5, the oscillogram being taken at the instant of current inrush.

Inspection of the oscillograms defi-

Fig. 3 (left). Rogowski coils in position on the duct

Fig. 4 (right). Portion of factory feeder duct under test

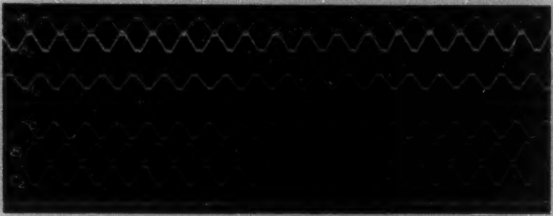


Fig. 5. Three-phase balanced load. Infrared lamp bank

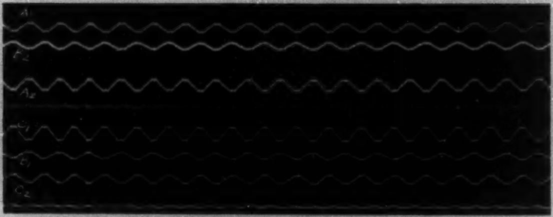


Fig. 6. Single phase load, infrared lamp bank, phase AC

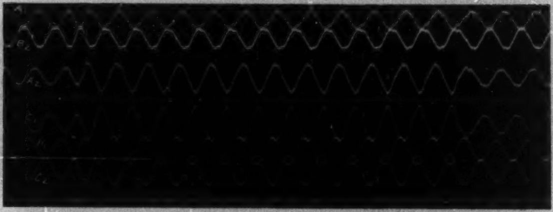


Fig. 7. Single-phase welder load added to phase AC of a balanced 3-phase infrared lamp load

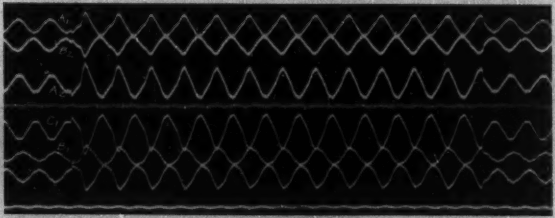


Fig. 8. Single-phase welder load added to a single-phase infrared lamp load, all on phase AC

nitely indicates the existence of the equal and opposite current characteristic for all load conditions. This can be noted by comparison of the paired bar current traces A_1 and B_2, A_2 and C_1, and B_1 and C_2 in all oscillograms in which the traces are readily shown to be about 180 degrees apart. The relationship exists whether the currents are balanced as in Fig. 5, unbalanced as in Fig. 7, or single-phase as in Figs. 6 and 8. A very interesting illustration of this phenomenon is the transient portion of infrared lamp load in Fig. 9.

The oscillograms also illustrate the tendency for even distribution of currents among all bus bars under various load conditions. For example, the addition of a single-phase load across phases *A*-*C* has caused an increase in the currents not only in the bars associated with these phases, but also in B_1 and B_2. The superposition of two signal-phase loads across the same phases, as in Fig. 8, is an even more striking example of this distribution tendency. The appearance of currents in bars B_1 and B_2 in the cases just referred to is, of course, the result of induction. This occurs because of an interphase tie made on the *B* bars. Quantitative results of current distribution and temperature rises with and without interphase ties are discussed in the following.

Laboratory Tests and Results

To study these phenomena further and to obtain data over a wider range, a series of laboratory tests were conducted under controlled load conditions. Three 10-foot sections of duct identical to the duct tested in the factory were used for this purpose. Rogowski coils were similarly employed for the simultaneous measurement of bar currents. Higher magnitudes and greater unbalance of currents were used so as to amplify differences for better analysis. A general view of the laboratory test setup is illustrated in Fig. 10.

The laboratory test series consisted of the following:

1. An approximate 3-phase balanced load.
2. A 3-phase load with some unbalance.
3. A 3-phase load with a greater unbalance than in test 2.
4. A single-phase load across phase *A*-*B* with an interphase tie on bars C_1 and C_2.
5. Similar to test 4 but with a lower value of phase current.
6. Similar to test 4 with the same phase current but with the interphase tie on bars C_1 and C_2 removed.

As with the factory tests, interphase ties were made on all phase bars at both the line and load ends except in test 6.

All measurements of bus bar currents, phase angles, and leg voltage drops were made directly from oscillograms taken for each test. Phase currents were obtained by the vectorial addition of the bar currents. Leg impedances were calculated from the phase currents and leg voltage drops. The results of these measurements and calculations are shown in Table I, where the current magnitudes and phase angles of paired bars are placed in adjacent rows for ready comparison. The circuit diagram of the laboratory tests, illustrated in Fig. 11, indicates

Fig. 9. Three-phase balanced infrared lamp load at instant of current inrush

points of voltage measurements and interphase ties at the beginning and end of the run.

Measurements of maximum and minimum bar temperatures and power losses were also made in single-phase tests 4 and 6. The same phase currents were maintained in these tests and the interphase connection at the load end of *C* phase was removed in test 6. The purpose was to determine the effect on temperature rise, power loss, and bus bar current distribution. The results of these tests are shown in Table II.

The first observation to be made from the test data is a general quantitative confirmation of the essentially equal and opposite current characteristics in paired bars for all tests, including test 6 where the only paired bars consisted of A_1 and B_2. It will be noted that there is some deviation in magnitude and phase angle from an exact equal and opposite relationship. In part, the deviation can be accounted for by experimental errors in recording and measurement. While the accuracy of both the phase-angle and current magnitude readings is affected by errors in measuring the oscillogram traces, the current magnitude values are further affected by the large turns ratio that must be employed with the Rogowski coils.

For the most part, however, the deviation is to be expected since it would be theoretically impossible to obtain exactly equal and opposite currents in a 3-phase configuration of three pairs of bars because of interpair proximity. The division of currents will approach this ideal only as the distances between pairs become infinite. From a practical standpoint, however, the relationship achieved in the paired-phase arrangement is remarkably close.

The significant result of equal and opposite currents in pairs of bars is the maximum neutralization of magnetic fields and its resultant negligible effect on magnetic housings or surroundings. Thus, no additional losses are obtained and reactance drops are kept low.

The tendency to divide currents evenly among bars, especially under unbalanced loads, is also to be noted from the test results. For instance, a decrease in the currents of bars A_1 and A_2 has occurred in tests 2 and 3 (unbalanced 3-phase) from test 1 (balanced 3-phase) even though the current in phase *A* has remained approximately the same for the three tests.

The significance of this tendency in a feeder system is to produce more even temperature rises among bars. Extreme temperatures in bars, having currents

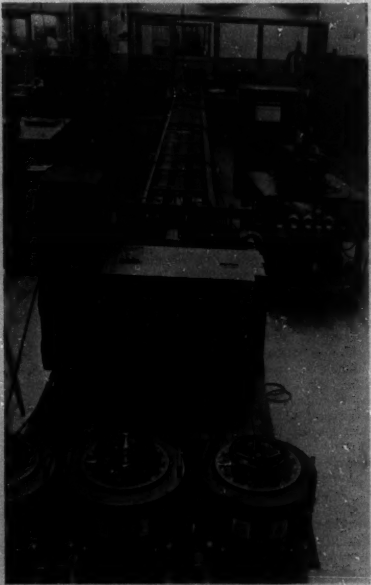


Fig. 10. Laboratory test setup

Table I. Laboratory Test Results

	1	2	3	4	5	6*
I_A, amperes	947	927	917	988	783	982
I_B	1,032	610	603	988	783	982
I_C	919	598	473			
I_{A_1}	542	485	470	600	470	694
I_{B_2}	539	448	456	648	511	773
ϕA_1-B_2, degrees	169	170	180	180	180	180
I_{A_2}	480	464	456	373	298	282
I_{C_1}	542	474	429	258	214	
ϕA_2-C_1	173	180	170	168	180	
I_{B_1}	578	251	195	356	286	244
I_{C_2}	508	197	108	258	204	
ϕB_1-C_2	176	165	140	171	169	
ϕA_1-A_2	46	23	10	0	0	22
ϕB_1-B_2	47	61	46	0	0	18
ϕC_1-C_2	59	60	63	180	172	
V_A, volts	0.368	0.369	0.372	0.361	0.288	0.459
V_B	0.323	0.196	0.196	0.311	0.252	0.356
V_C	0.375	0.208	0.173			
Z_A, ohms $\times 10^{-6}$	388	398	406	366	368	468
Z_B	313	322	325	316	322	363
Z_C	411	348	366			

* Interphase connection removed at load end of *C* phase.

Table II. Power Loss and Temperature Data for Single-Phase Tests

Test No.	I, Amperes	W, Watts	Maximum Bar Temperature, °F	Minimum Bar Temperature, °F	Spread, °F	Maximum Current for 185 Degrees† Operating Temperatures
4	988	534	118	89	29	1,700
6*	982	613	135	83	52	1,350

* Interphase connection removed at load end of *C* phase.

† Duct of this type is normally rated with this as a maximum operating temperature, 99 degrees Fahrenheit (F) (55C) rise above an 86F (30C) ambient = 185F (85C).

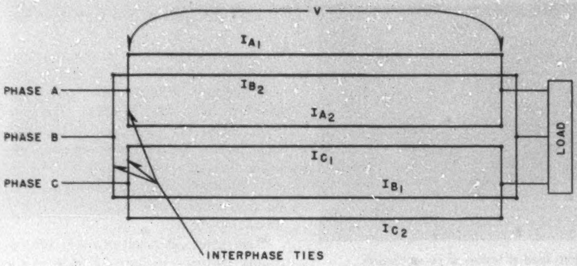


Fig. 11. Circuit diagram of laboratory tests

otherwise distributed, are minimized. As a result, efficient use is made of the conductor material of all bars and a cooler operating feeder is obtained for practically all service conditions. Other important advantages to be gained by the equal division of currents among bus bars include a wider safety margin for overload conditions, and a reduction of hot spots in bars and associated equipment tied to the duct.

The importance of interphase tie connections at load points is demonstrated by the results of single-phase tests 4 and 6 shown in Tables I and II. The same phase currents were used in both tests. Lack of a tie of the *C* phase bars in test 6 disturbed the impedance of the system and resulted in higher currents in the paired bars A_1 and B_2, and lower currents in the isolated and unpaired bars B_1 and A_2. Although the equal and opposite current characteristic was maintained in the paired bars A_1 and B_2, the increased spread in current distribution among all bars materially increased both the maximum bar temperature and the difference between maximum-minimum bar temperatures. Voltage drop and watts loss also increased. Of interest also is the decrease in current-carrying capacity for a given temperature rise [55 degrees centigrade (C)] from 1,700 to 1,350 amperes.

It may be obvious that interphase ties are necessary in any multibar feeder system if it is to perform efficiently. This importance, however, should not be overlooked in the original design of the configuration of bus bars. The ease with which the ties can be made in the field will assure that they will be used and that faulty or improper connections will be eliminated.

As a final point of observation, it is to be noted that, regardless of current distribution among the bars, voltage drops and phase currents remained proportional. For example, tests 1 and 2 show similar currents for phase *A* while the currents in the two bus bars A_1 and A_2 are different because of changes in the total load. However, the voltage drops in phase *A* for these tests are approximately the same. From the experimental data, therefore, it would appear that for a given phase current, the phase voltage drop is constant regardless of the current distribution established in the bus bars of that phase by other phase currents.

Conclusions

1. In a paired-phase multiple bus bar arrangement, the currents in paired bars distribute themselves essentially equal in magnitude and opposite in phase under all conditions of balanced, unbalanced, transient, and steady-state loads. Minimized or almost complete cancellation of magnetic fields is therefore obtained.

2. A tendency exists for phase currents to equalize themselves among bus bars for conditions of unbalanced load. A consequent decrease and more uniform temperature rise among the bars, elimination of hot spots, and a wider safety margin for overload conditions is obtained.

3. The use of interphase ties at load connecting points in a feeder system is necessary for obtaining a more uniform division of current among all bars, low wattage loss, and more uniform temperature rises in the bars.

References

1. Paired-Phase Bus Bars for Large Polyphase Circuits, **Lawrence E. Fisher, Robert L. Frank.** *AIEE Transactions (Electrical Engineering)*, vol. 62, Feb. 1943, pp. 71–77.

2. Instrument Transformers (book), **B. Hague.** Isaac Pitman and Sons, Ltd., London, England, Chap. 12, 1936.

3. Design and Behavior of Magnetic Potentiometers, **J. Ernest Shaw.** *The Electrical Review*, London, England, vol. 102, Apr. 20, 1928, pp. 678–80.

◆

Discussion

Lawrence E. Fisher (General Electric Company, Plainville, Conn): The authors state that it is to be expected that the need for interphase ties at load connection points is important regardless of the bus bar configuration since electrical omission of any bar would have an effect on the impedance of the system.

In the interlaced arrangement used by my company, no difficulty is experienced because of unbalanced 3-phase loads or single-phase loads. It is not necessary to tie the *C* phase bars in parallel at a single-phase *A-B* tap off.

Therefore, it is believed that the third item in the "Conclusions" is incorrect unless it is reworded to indicate that it applies only to the paired-phase busway. Having originated the paired-phase principle, I am familiar with the peculiarities of this busway as brought out in the paper. It was known that single-phase loads (or unbalanced 3-phase loads) caused undesirable voltage drops and unequal division of current between the two bus bars in parallel, unless both bars of all three phases were tied together at the tap-off point. Engineering instructions were issued to tie both bars together in all three phases at every tap off even for a single-phase tap.

A helpful way to understand this peculiar phenomenon in the paired-phase busway is to think of the pairs as being transformers. For example, Fig. 1 shows if we assume a single-phase *A-B* load with the *C* bars not connected together at the load, the A_1 and B_2 bars will carry more than two-thirds of the total current because of the wider spacing of the A_2 and B_1 bars.

Now both the A_2C_1 pair and the B_1-C_2 pair may be considered as transformers even though there is no iron core. It was proved by test that the current flowing in A_2, for example, would cause a current to flow in C_1 which would lag the current in A_2 by about 180 degrees and, also, it would lag the current induced in the C_2 bar by 180 degrees. Consequently, placing connecting tie straps between the *C* bars at both ends of the run effectively short-circuits both of these would-be transformers. Now the result of short-circuiting the secondary of a transformer is a reduction in the impedance measured across the primary terminals. Since the primaries of these transformers are the A_2 and B_1 bars, it is anticipated that the impedance of these bars would be reduced by adding the tie straps between the *C* bars. This is, of course, exactly what happens to the extent that the total single-phase current is then divided more evenly between the A_1 and A_2 bars and the B_1 and B_2 bars.

For the interlaced arrangement in which the bars are arranged *ABCABCABC*, the total current divides fairly evenly between the bars in parallel with or without the straps connecting the *C* bars in parallel. No appreciable voltage is induced in the *C* bars from one end to the other.

Thomas J. Higgins (University of Wisconsin, Madison, Wis.): I find this a very clearly written paper. Without question, the experimental data advanced substantiate the statements in the author's conclusions. This paper complements very nicely the paper by Fisher and Frank (ref. 1 of the paper), wherein attention was confined, in the large, to steady-state balanced load conditions.

J. B. Cataldo and N. Shackman: The authors appreciate the interest and comments offered by the discussers.

Mr. Fisher comments that our conclusion of the need for interphase ties should be restricted to the paired-phase arrangement. Our paper was limited to the paired-phase arrangement and we have not attempted to measure actually the effect of interphase ties on interlaced bus bar arrangements such as those used by Mr. Fisher's company.

Theoretically, we all know that the existence of a short-circuited secondary loop in close proximity to a primary loop will result in a reduction of impedance measured across the primary loop. These secondary loops exist in any multibar feeder system and the extent to which the bus impedance is reduced will depend, of course, on other factors such as the relative position of the bus bars. One of the purposes of our paper was to test this theory by actual measurements.

We believe that it would be a contribution to the general knowledge of the art and, therefore, most desirable if Mr. Fisher would present data giving actual measurements of the effects on the interlaced bus bar arrangements to which he refers in his comments.

Correlation of Rating Data for Rotating Electric Machinery

ERWIN O. A. NAUMANN
NONMEMBER AIEE

THE variety and extremes of environmental conditions to which electric equipment is subjected in military aircraft in connection with the stringent performance, weight, and space requirements make it mandatory that the environmental factors be properly considered in the application of this equipment. The major effect of the environment is that of cooling which has become a problem area in high-speed, high-altitude aircraft.[1–3] For some time, the available information on cooling characteristics of electric machinery has been inadequate in its scope and was not sufficiently confirmed by tests.[4–6] Adequate and correct rating data must be made available to the airframe designer for the purpose of applying the equipment in an efficient and reliable manner. In this paper, a rating and correlation method is described as applied to blast-cooled aircraft generators, but the method is not restricted to this equipment and is adaptable to any equipment which is cooled by forced convection with gases or liquids as the coolant.

Nomenclature

A_f = flow area
A_s = surface area
c_p = specific heat
C, C_1, C_2 = constants
g = acceleration of gravity
h = heat-transfer coefficient
κ = ratio of specific heats
K = constants
L = electrical load
m, n = exponential constants
N = speed
p = pressure
p_a = static outlet (ambient) pressure
ΔP = pressure drop
$\overline{\Delta P}$ = corrected pressure drop
Q = heat flow
R = gas constant
t = temperature
t_i = inlet temperature
T = absolute temperature
T_i = total inlet temperature
T_e = total outlet temperature
T_0 = reference temperature
T_w = wall or surface temperature
W = weight flow
$\overline{W}$ = corrected weight flow
$\alpha, \beta, \gamma, \delta$ = exponential constants

Theory

The purpose of the correlation method is to develop complete relationships between the essential environmental factors and the allowable output of the equipment with a minimum amount of test work. Like prior studies,[7–10] a basic theory is developed with simplified assumptions; however, the only purpose of the theoretical considerations is the development of useful correlation factors. The final result of the rating method is dependent on adequate and properly correlated test data. The correlation takes into regard three basic factors: the air-flow characteristics, the heat-transfer characteristics, and the machine operating characteristics.

Air-Flow Characteristics

Flow phenomena by forced convection are governed by Mach number, Reynolds number, and pressure ratios. With heat addition, the enthalpy ratio is an additional factor. This relationship holds not only for simple conditions but also for flow systems of any complexity as long as all independent variables are properly represented. The Mach number is a velocity divided by a sound velocity, the Reynolds number is a flow velocity times density divided by a viscosity, and the enthalpy ratio is a temperature ratio. The flow Mach number can be represented by the factor $W\sqrt{T}/p$, since the sound velocity is proportional to the square root of the temperature, and the speed Mach number by the factor $N/\sqrt{T}$, since the circumferential speed of a rotor is proportional to the revolutions per minute; see "Nomenclature" for definitions. The factor representing the Reynolds number is $W/T^{0.76}$, because the viscosity is approximately proportional to $T^{0.76}$. Theoretically, it is immaterial to what specific locations in a system these factors pertain. However, since the complete flow system is being considered, the inlet and outlet conditions apply with regard to pressures and temperatures. Also, it is a matter of choice whether to use total or static values of pressures and temperatures. The proper selection depends on which values are most useful with regard to the application and to a simple form of correlation. For forced air flow through generators the following values directly apply: pressure drop (total inlet pressure minus static outlet pressure), static outlet pressure, and total air-inlet and outlet temperatures. Therefore, the complete correlation of flow conditions for a specific system is accomplished if all of the following factors are included:

Flow Mach number: $W\sqrt{T_i}/p_a$
Reynolds number: $W/T_i^{0.76}$
Pressure ratio: $\Delta P/p_a$
Speed Mach number: $N/\sqrt{T_i}$
Enthalpy ratio: T_e/T_i

Evaluation on the basis of all of the foregoing factors would be a tedious procedure. For this reason, a simplified

---

Paper 54-355, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 10, 1954; made available for printing August 4, 1954.

Erwin O. A. Naumann is with the Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

procedure is taken based on the following considerations. If as a first approximation, the system is equivalent to that of an orifice or duct of constant cross section, the following relationships hold, if ΔP and p_a are of identical dimensions

$$\frac{1}{A_f}\sqrt{\frac{R}{g}}\,\frac{W}{p_a}\sqrt{T_t}=\sqrt{\frac{2\kappa}{\kappa-1}\left(1+\frac{\Delta P}{p_a}\right)^{\frac{\kappa-1}{\kappa}}\left[\left(1+\frac{\Delta P}{p_a}\right)^{\frac{\kappa-1}{\kappa}}-1\right]} \tag{1}$$

for $\frac{\Delta P}{p_a}\leq\left(\frac{\kappa+1}{2}\right)^{\frac{\kappa}{\kappa-1}}-1$

$$\frac{1}{A_f}\sqrt{\frac{R}{g}}\,\frac{W}{p_a}\sqrt{T_t}=\sqrt{\kappa\left(\frac{2}{\kappa+1}\right)^{\frac{\kappa+1}{\kappa-1}}}\left(1+\frac{\Delta P}{p_a}\right) \tag{1A}$$

for $\frac{\Delta P}{p_a}\geq\left(\frac{\kappa+1}{2}\right)^{\frac{\kappa}{\kappa-1}}-1$

It is found that these equations can be simplified with a minor error which is less than 3.3 per cent to

$$\frac{1}{A_f}\sqrt{\frac{R}{g}}\,\frac{W}{p_a}\sqrt{T_t}=\sqrt{2\frac{\Delta P}{p_a}} \tag{2}$$

for $\frac{\Delta P}{p_a}\leq 1.0$

$$\frac{1}{A_f}\sqrt{\frac{R}{g}}\,\frac{W}{p_a}\sqrt{T_t}=\frac{1}{\sqrt{2}}\left(1+\frac{\Delta P}{p_a}\right) \tag{2A}$$

for $\frac{\Delta P}{p_a}\geq 1.0$

Therefore, with mass flow less than critical, the following simple relationship holds with C being a systems constant

$$W=\frac{C}{\sqrt{T_t}}\sqrt{p_a\Delta P} \tag{3}$$

Plotting test data for log $(p_a\Delta P)$ versus log W should, for constant total air-inlet temperature, show a straight line with a slope of 2. For critical flow, i.e., $\Delta P/p_a>1.0$, the equation changes to

$$W=\frac{C}{2\sqrt{T_t}}(p_a+\Delta P) \tag{3A}$$

with C being the same constant as in equation 3.

To account for pressure losses caused by turbulence and friction, they are normally expressed as a resistance factor which, multiplied by the velocity head, gives the pressure drop encountered. This resistance factor can be a constant or a function of the applicable Reynolds numbers, e.g., for turbulent flow in ducts it is inversely proportional to the fourth root of the Reynolds number. Since the velocity head is proportional to the product of static pressure times the square of the previously defined flow Mach number, the interrelation between pressure drop caused by friction, flow Mach number, and Reynolds number is of the type

$$\frac{\Delta P}{p_a}\propto\left(\frac{W}{p_a}\sqrt{T_t}\right)^2\Bigg/\left(\frac{W}{T^{0.75}}\right)^{\alpha} \tag{4}$$

with α from 0 to 0.25.

This equation is similar to the one for orifice restriction with the exception that the mass-flow exponent may differ from 2, i.e., between 1.75 and 2. Combining both effects, i.e., the pressure drop caused by flow restriction in a duct system with the pressure drop caused by friction and turbulence should result in this general type of equation

$$W=\frac{C}{T_t^{\beta}}(p_a\Delta P)^{1/\gamma} \tag{5}$$

with β from 0.68 to 0.50 and γ from 1.75 to 2.0. An additional loss may be encountered because of heat addition to the air. For example, the pressure drop relationship for heat addition in a constant area duct reads

$$\frac{\Delta P}{p_a}\propto\left(\frac{W}{p_a}\sqrt{T_t}\right)^2\left(\frac{T_e}{T_t}-1\right) \tag{6}$$

Since this equation is of the same type previously established, the two can be combined by considering a mean in lieu of the inlet temperature, resulting in

$$W=\frac{C}{(T/T_0)^{\beta}}(p_a\Delta P)^{1/\gamma} \tag{7}$$

with β from 0.68 to 0.50, γ from 1.75 to 2.0, and T from T_t to T_e. β and γ are systems constants. The function which correlates T to T_t and T_e is also a systems constant. An additional effect which must be considered is that of the speed of the rotor or fan. For an approximate correlation, the induced pressure head is proportional to a certain power of the velocity head taken as function of the circumferential speed of the rotor and is independent of the Mach number

$$\frac{\Delta P}{p}\propto\left(\frac{N}{\sqrt{T}}\right)^{\delta} \tag{8}$$

with δ close to 2.0. This effect may be positive or negative, thus either adding or subtracting from the external pressure drop across the system. Since this effect is minor, at least for blast-cooled generators, it will probably not introduce an essential error if the variation in temperature is neglected. This results in the proposed air-flow correlation equation for the complete system

$$W=\frac{C}{(T/T_o)^{\beta}}[p_a\Delta P+C_1(p_aN)^{\delta}]^{1/\gamma} \tag{9}$$

If the flow Mach number is of considerable influence on the speed effect, this correlation equation will be affected by a change in the numerical values of the constants. C and C_1 are systems constants and C_1 can be positive or negative. The exponents β, γ, and δ are also systems constants within the following probable ranges: β from 0.68 to 0.5, γ from 1.75 to 2.0, and δ close to 2.0. T is a mean value between T_t and T_e and as such a constant function of these two temperatures. T_o is an arbitrarily assigned reference temperature. The recommended procedure for determining these systems constants is to use the test data of the later described rating tests which determine minimum air flow and pressure drop versus air-inlet temperature and electric output at the maximum permissible machine temperatures. In spite of relatively large variations in air-inlet temperature, the air-outlet temperature will vary only slightly or, in case the air-outlet temperature represents the critical limit, will not vary at all. The thus obtained data of W and p_a (ΔP) shall in the first approximation be plotted on log-log paper, i.e., as log $[p_a(\Delta P)]$ versus log W. A straight line which averages these test points will determine the value of γ which is the slope of the drawn line.

The second step involves the possible temperature change correction. If it is obvious from this plot that the mass flow test data at the lower air-inlet temperatures are consistently larger than those at the higher air temperatures and if this effect appears essential, straight lines should be drawn parallel to the originally drawn line which average the test data at each air-inlet temperature. The deviation of these lines from the original will then determine a correction factor of the type $(T/T_o)^{\beta}$ with an arbitrarily assigned reference temperature T_o and with T a mean value of T_t and T_e such as $1/2\,(T_t+T_e)$ or $\sqrt{T_tT_e}$ or any suitable function. If T_e is constant or nearly constant, T is a function of T_t alone.

The third step is a replot of the data showing log $[p_a(\Delta P)]$ versus log $W(T/T_o)^{\beta}$ to determine the possible speed effect correction. If such an effect is essential, it will show up in large deviations of the test data from the straight line at low values of $\Delta P/p_a$, i.e., low air flows at high ambient pressures. To determine the correction factor $C_1(p_aN)^{\delta}$, it is practical to take two of the most extreme deviations from which the two unknown constants C_1 and δ can be calculated. Applying the thus obtained correction factors to the other data will then show whether minor changes in C_1 and δ must be made to obtain better correlation.

A further consideration is the determination of the choking condition which is the condition where the air flow is only a function of the total inlet pressure, i.e., $p_a+\Delta P$, regardless of the ambient or outlet pressure p_a. This choking condition is represented by the critical value of $\Delta P/p_a$ which from the prior analysis for nozzle flow is about unity. Because of the effects of flow resistance and heat addition, the critical value will be higher. This is also true if the system is not representative of a single nozzle but of a series of nozzles, but the speed effect can cause the critical value to be higher or lower. Using the symbol K for the critical value of $\Delta P/p_a$, with ΔP and p_a being of the same dimension, the air-flow correlation equation changes to

$$W=\frac{C}{(T/T_o)^{\beta}}\left[\frac{\sqrt{K}}{1+K}(p_a+\Delta P)\right]^{2/\gamma} \tag{9A}$$

for $(\Delta P)/p_a>K$. K is a systems constant and will not be smaller than the value of 1.0, since the rotative speed effect on the critical pressure ratio will have less effect than the other factors for the type of equipment considered. The results of the complete air-flow correlation including choking conditions, can then be shown in graph form with log p_a as abscissa and log $\overline{W}$ as ordinate with $\overline{\Delta P}$ as parameter

$$\overline{W}=C(p_a\overline{\Delta P})^{1/\gamma} \tag{10}$$

for $\dfrac{\Delta P}{p_a}\leqq K$

$$\overline{W}=C\left[\frac{\sqrt{K}}{1+K}(p_a+\overline{\Delta P})\right]^{2/\gamma} \tag{10A}$$

for $\dfrac{\Delta P}{p_a}\geq K$

The correction factors, if applicable and necessary, are as follows

$$\left(\frac{T}{T_o}\right)^{\beta}=\frac{\overline{W}}{W}$$

and

$$C_1(p_aN)^{\delta}=p_a(\overline{\Delta P}-\Delta P)$$

These can be shown in a separate plot or tabulation or directly on the graph as deviations, e.g., if only two or three speeds need to be shown.

HEAT-TRANSPER CHARACTERISTICS

Heat-transfer phenomena by forced convection are governed by the Nusselt number, Reynolds number, Prandtl number, and Mach number. The Prandtl number is a material constant. Its slight variation with temperature can be neglected. The Mach number has for subsonic velocities no influence on the Nusselt number other than its effects on the adiabatic wall temperature, which is the temperature the surface would assume in the gas stream without heat transfer, and on the reference temperature for the determination of property values like viscosity, conductivity, etc. Since supersonic velocities are not involved, the practical way to eliminate the Mach number effect on local heat transfer is to relate all temperatures to total temperatures in lieu of static temperatures in the air stream. For local-heat transfer conditions, the assumption of the Nusselt number being a sole function of the Reynolds number is warranted. The difficulty is the translation of local conditions to correlation data which are valid considering the equipment as a system. It appears that the Mach number and rotative speed effects will play a role since they govern the changing flow conditions through the different passages. Further, the heat gain of the air at different locations will not only affect the temperature potential between surface and the cooling air but also will influence the flow distribution. Also, the location and magnitude of heat sources in the machine will vary with the electric load conditions and will not be a simple function of the total losses. In addition, the assumption of forced convection may not be entirely applicable inasmuch as natural convection effects, especially on the exterior surfaces of the equipment, and conduction effects through the pad or mounts are at least to some degree of influence.

Because of these complications, which are not susceptible to a rigorous analysis, a necessarily simplified procedure must be chosen, but precautions must be taken not to draw unwarranted conclusions from a simplified analysis without experimental proof. The Nusselt number is a ratio of heat-transfer coefficient to thermal conductivity, the latter being approximately proportional to the temperature in the 0.75 power. For simple flow conditions in the turbulent region along ducts and plates, the Nusselt number is proportional to the Reynolds number with a power coefficient of 0.75 to 0.80. For laminary flow the power coefficient approaches 0. In general, the following relationship holds

$$\frac{h}{T^{0.75}}\propto\left(\frac{W}{T^{0.75}}\right)^{m} \tag{11}$$

(m within 0.50 to 0.80) or neglecting the small influence of the absolute temperature

$$h\propto W^{m}$$

Assuming heat transfer in a duct of constant cross section with constant surface temperature T_w, the general equation for the heat flow reads[11]

$$Q=hA_s\frac{T_e-T_i}{\ln\dfrac{T_w-T_i}{T_w-T_e}} \tag{12}$$

To obtain Q as a function of W, T_i, and T_w, the variables T_e and h must be eliminated by use of the following equations

$$Q=c_pW(T_e-T_i)$$

$$h\propto W^{m}$$

It can be shown analytically that in lieu of using this complicated function, a very good approximation is

$$Q\propto W^{n}(T_w-T_i) \tag{13}$$

where $m<n<1.0$.

The effect of the heat gain of the cooling air can thus be incorporated by increasing the air-weight flow exponent but still referring the temperature potential to the air-inlet temperature. Others have arrived at similar relationships.[7,9,10,12] As a basis for the heat-transfer correlation, an equation of the following type should suffice

$$Q=C_2W^{n}(T_w-T_i) \tag{14}$$

This equation is also applicable in case the air-outlet temperature becomes limiting. Then T_w is identical to T_e and the exponent n is 1.0, with C_2 being the specific heat of air. In general, however, T_w in the equation need not necessarily be identical to a hot-spot temperature but can be considered a machine constant which, when properly determined, assures that the applicable hot-spot temperature will not be exceeded if the equation is satisfied. It can be called the "equivalent" wall temperature.

If there were definite assurance that, disregarding the differences in the designs, the heat-transfer relationship of all blast-cooled machines would stringently follow an equation of the foregoing type and if T_w were either identical to the applicable measurable hot-spot temperature or a known sole function of it, the determination of the constant C_2 and n would be a simple matter since the absolute magnitude of T_w and T_i as well as the pressure level are of no influence. In this case, tests at different altitudes and air-inlet temperatures would be unnecessary. Work along these lines is now in progress and preliminary favorable results have been reported.[13] However, in view of the expressed concern of oversimplifying the problem, which at least at the present time may cast doubt on the dependability of results, tests are proposed at stabilized limiting machine temperature conditions within the possible application ranges of temperatures and pressures. Essentially the test method is designed to develop the heat-transfer relation which envelops all possible conditions by interpolation of test data. Experience with this method may eventually prove that a simpler method would be satisfactory. In detail, the following procedure is proposed:

At a number of constant load conditions, the air-inlet temperature and the air outlet pressure should be varied in a number of steps, like 40 degrees centigrade (C) in air-inlet temperature and pressures equivalent to 20,000 foot-steps. For each condition the air flow should be so adjusted that one of the selected and measurable critical component temperatures, when stabilized, reaches its limiting value. With Q denoting the total losses of the machine and W the measured minimum permissible flow, the values of the constants C_2, n, and T_w have to be analytically determined in such a manner that the ratio of $C_2W^n(T_w-T_i)$ over Q is either equal or smaller than unity but approaches the value of 1.0 as closely as possible for all test conditions where one and the same critical component is limiting.

A practical method of determining the constants is to assume T_w to be equal to the actual temperature limit. Then $\log(Q/T_w-T_i)$ is to be plotted against $\log W$. The values of C_2 and n can be easily determined by means of a straight line drawn through the test points. Marking the different air-inlet temperatures will then show whether the assumed value of T_w should be corrected for better correlation. Scrutiny of the test data with regard to the different altitude conditions will also show whether the incorporation of an altitude correction factor would be necessary. If more than one of the critical components reaches a limiting temperature, either this analytical procedure can be applied separately for each component or an attempt can be made to correlate all test data by a single equation regardless of the dislocation of the hot spot. The results obtained from this heat-transfer correlation can be shown in graph form with $\log Q$ as abscissa and $\log W$ as ordinate and with T_i as parameter. The resulting curves are straight lines. If more than one of these equations have been determined, only that portion of each equation shall be shown which gives the lower heat-dissipation rate and W and T_i being equal. In this case, each temperature line will consist of two or more intersecting straight lines.

MACHINE OPERATING CHARACTERISTICS

The correlation of electric output rating with the total losses must include the controllable essential factors such as speed and power factor. Considering the types of losses, they can be put into the following general categories: losses dependent on speed such as brush and bearing friction, and windage; losses dependent on voltage such as core losses; and losses dependent on output current such as copper and stray losses, brush drop losses, and exciter losses. The latter can be broken down into the same general categories. Since voltages are to be held within close limits, the total losses can be separated into speed losses and load losses. The load losses are mainly electrical resistance losses and as such are proportional to the square of the output current. No general function of the speed losses can be assumed. However, for the type of equipment considered, the magnitude of the speed losses is within 15 to 40 per cent of the total losses at rated output conditions. The machine temperature pattern will probably affect the total heat loss only to a minor degree since the losses shall be taken at limiting machine temperature conditions; however, the environmental effects on bearing and brush friction as well as brush drops appear rather unpredictable.

The practical approach to develop the required relationship is to use the efficiency test data, the manufacturer's design data analyzing the losses of the machine, and the loss data obtained from the already described thermal rating tests. Thus, data taken from torque-speed, electrical and air-flow temperature rise measurements are to be used. These data should be plotted as output current against $\log Q$ with speed and power factor as parameter, whichever applicable. In lieu of output current, power output in kilowatts or kilovolt-amperes can be used. This plot has been found useful since the resulting curves are nearly straight lines which can be analytically proved if the assumption is correct that the total losses consist of speed and load losses.

RATING CHART

On a rating chart, all three of the discussed characteristics can be put together by aligning the air-flow chart with the heat-transfer chart and the latter with the machine characteristics (loss-load) chart, thus combining all essential factors for proper application for the equipment in the customary dimensions. The advantage of this presentation is that, in addition to the environmental and operating data, the air flow and losses can be read to enable the airframe builder to make a performance and heat-load analysis for the cooling system. The data on this chart refer to stabilized thermal conditions at maximum permissible machine temperatures. The need of providing rating information for transient thermal conditions will become more acute. Preliminary theoretical investigations indicate that the inclusion of transient characteristics will not make the rating chart useless or obsolete since it can be effected by a supplementary plot.

Rating Tests

Tests were conducted at the Equipment Laboratory, Wright Air Development Center, to prove the recommended correlation method outlined in the previous section. Only the important aspects required to develop a complete understanding of the methods involved will be illustrated in this paper; further detailed information can be found in the official records.[13,14]

GENERAL CONSIDERATIONS

Since the purpose of the correlation method is to determine from a minimum number of test data all essential factors for proper rating of the equipment, it is obvious that the results of the method can only be as good as the quality and accuracy of the test data. Therefore, the development of suitable test techniques and the use of adequate measuring and recording devices are of great importance. Judicious selection of test conditions will reduce the number of tests and minimize the necessity of extrapolation.

INSTRUMENTATION

The instrumentation to provide the necessary data can be classed in three general categories: devices for measuring temperatures, pressures, and electric load. Air-flow and heat dissipation data are usually calculated from pressure and temperature measurements at constant area restrictions and temperature rises, although variable air-flow meters (rotameters) and power measuring devices are also used. Industry standards cover the basic principles and techniques. Their implementation to the basic rating test procedure requires special consideration because of the principle involved of measuring temperatures of critical components and limiting these temperatures to some assigned value.

In view of the dependency of the design, the materials, and the manufacturing procedures on the location of hot spots and their temperature limits, the proper instrumentation or information must be provided by the designer. Special measuring and protective devices are being developed;[15,16] one was used in the later described test on a 300-ampere d-c generator. Because of the practical difficulties in attaching, monitoring, and maintaining these devices which otherwise are indispensable for scientific studies,[13] the determination of mean winding temperatures through resistance measurements may be considered adequate because tests which are not yet completed have shown so far that hot-spot temperatures can, with a certain margin of error, be related to the mean winding temperature. Another important problem is the adequate determination of the air-outlet temperature because of the nonuniform flow pattern at or around the air-outlet openings.

TEST TECHNIQUES

The employed technique requires that load, air-inlet and ambient temperature, and pressure altitude are held constant and that the pressure head be so regulated that any of the critical temperature limits are approached by 4 C and remain stabilized within 2 C for a 10-minute period. The first attempts to follow this technique necessitated chamber test operation of up to 2 hours' to obtain the desired stabilized conditions. With further experience, the average time for obtaining satisfactory data readings was reduced to 60 minutes. This time included the change in ambient pressure conditions of the chamber. Since change in temperature conditions required additional time because of the large thermal lag of the chamber, all tests at one air-inlet and ambient temperature were conducted consecutively to hold the total test time to a minimum.

SELECTION OF TEST CONDITIONS

For proper selection of test conditions, the following factors should be taken into consideration: speed range and output range of the equipment; and temperature, altitude, and pressure drop range conditions which can be expected in aircraft applications. An estimate of the latter conditions is reflected in recommended changes to present specifications which include environmental ranges as follows: air-inlet and ambient temperatures: −60 to 150 C; air-outlet pressures: 15 to 0.8 pounds per square inch (sea level to 65,000 feet); and pressure drops up to 400 inches of water. However, for the present type of equipment and the present applications, the air-inlet temperature and pressure drop extremes generally do not exceed 80 C and 50 inches of water respectively.

To cover adequately these ranges, tests are required at four to five different temperatures such as in steps of 40 C, and at three to four different air-outlet pressures such as in steps equivalent to a 20,000-foot altitude. Thus, 12 to 20 tests would completely cover any combination of environmental conditions. To repeat all these tests even for a small number of different load, load speed, or load-power factor conditions would be costly and time-consuming. For this reason, complete environmental testing is recommended for only one load condition, preferably the rated load condition, with testing at other loads and speeds or power factors conducted at only two different, but extreme, environmental conditions, such as high air-temperature–sea-level and low air-temperature 60,000-foot conditions. Taking as an example that three speed and three or four load conditions (0, 50-per-cent, 100-per-cent, and 125-per-cent rated load) are selected, the combination of these amounts to 8 or 11 respectively in addition to the rated load condition. Therefore, with a total number of tests of 20 to 40, a complete rating investigation can generally be accomplished. In many instances, limitations of the test chamber facilities will enter the considerations of selecting the test conditions and may necessitate extrapolation of test data to some degree. Further, since sea level test conditions do not require chamber facilities but only a conditioned air supply the actual altitude chamber test time can be reduced by conducting the sea level tests on a drive stand.

RESULTS OF TESTS OF A 40-KVA GENERATOR

The prototype of a 40-kva generator manufactured by the Westinghouse Electric Corporation, type no. *8QL40G*, was subjected to the rating tests in accordance with the outlined procedures. This machine has window-type air-inlet openings; therefore, the total pressure drop refers to the difference between the total pressure in the space around the air entrance windows and the static outlet pressure. The thermal limits of the machine were attained when the mean alternator field temperature, which was determined by resistance measurement, reached 185 C or the exciter shunt temperature, which was determined by thermocouple measurement, reached 145 C. Temperatures of other components were also measured, such as bearings, brushes, stator, and frame, but were not found limiting at any of the test conditions. The range of test conditions was limited by the available altitude chamber and drive stand facilities, and included the following values: Air-inlet temperature form −40 to 70 C; altitude from sea level to 60,000 feet; pressure drop form 0.3 to 10 inches of water; and output rating from no load to 100-per-cent load at 0.75 power factor. The ratio of pressure drop to outlet pressure, both in identical dimensions, ranged from 6×10^{-4} to 0.18.

Eighteen different test conditions were investigated and nine additional tests were conducted to determine the accuracy and degree of reproducibility. Twelve tests were conducted with rated load and rated speed; five at half-rated load and one at no-load conditions. The air-flow correlation curve giving log $(p_a\Delta P)$ versus log W is shown in Fig. 1. The different symbols refer to air-inlet temperatures. Adequate correlation was obtained by the equation

$$W=2.3\sqrt{p_a\Delta P} \tag{15}$$

where

W = air flow, pounds per minute
p_a = outlet pressure, pounds per square inch
ΔP = pressure drop, inches of water

It was unnecessary to incorporate speed and temperature correction factors. This equation was further substantiated by air-flow measurements with a rotameter under sea level conditions. Although a change in hot-spot location from the alternator field to the exciter shunt was observed, satisfactory correlation was obtained by a single equation, as illustrated in Fig. 2, assuming an "equivalent wall temperature" of 145 C which is equal to the lower of the established thermal limits. The applicable equation is

$$Q=0.5\times W^{0.78}(145-t_i) \tag{16}$$

where

Q = heat rejection rate, British thermal units per minute (Btu/min)
W = air flow, pounds per minute
t_i = air-inlet temperature, C

A plot of load and heat losses is shown in Fig. 3. The crosses denote the calculated data by the manufacturer. A considerable spread of test data was observed which are believed to be caused by high heat transfer from the surface of the machine because of the special test set-up. The dashed line assumes a fixed loss of 102 Btu/min and a loss variation proportionate to the square of the load. Because of the small amount of data at part

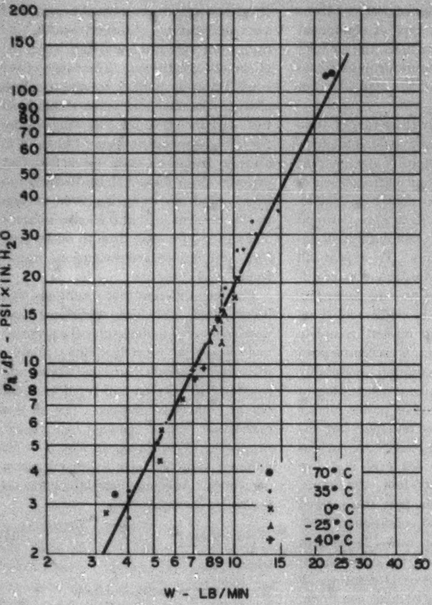


Fig. 1 (above). Air-flow correlation, a-c generator

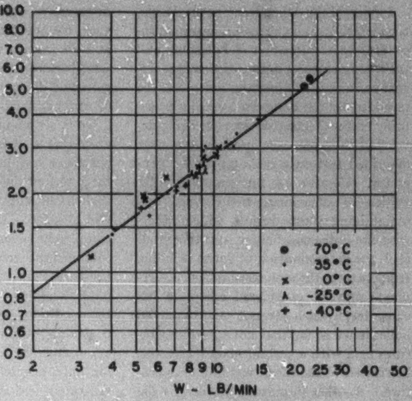


Fig. 2. Heat-transfer correlation, a-c generator

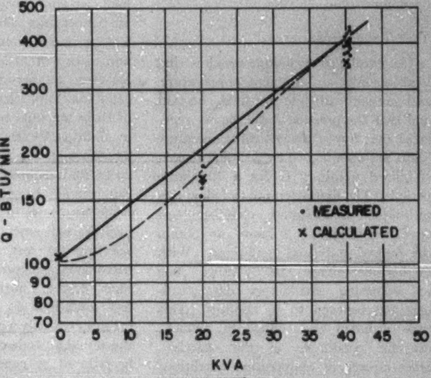


Fig. 3 (right). Loss-load correlation, a-c generator

load, the more conservative correlation equation was chosen which is depicted by the solid line

$$Q=102.5\times 4^{(L/40)} \qquad (17)$$

where

Q = heat rejection rate, Btu/min
L = load in kva at 0.75 power factor and 6,000 rpm

Fig. 4 is the resulting rating chart on a reduced scale. A typical example is illustrated but the use of the chart is not limited to this sequence. Taking the conditions of an air-outlet pressure equivalent to a 40,000-foot altitude (2.72 pounds per square inch), a pressure drop of 4.6 inches of water, and an air-inlet temperature of −25 C, then by following the arrows, these data are obtained: per-cent rated air flow is 33.5 per cent, per-cent heat rejection is 100 per cent, and per-cent rated output is 100 per cent. Multiplying the percentage figures with the reference data listed on the chart results in an air flow of 8.15 pounds per minute, heat rejection rate of 410 Btu/min which is equivalent to 7.2 kw, and allowable electric load of 40 kva at 0.75 power factor.

Assuming all test data being of equal merit, an estimate of the accuracy of the chart can be obtained by averaging the absolute deviations of the test data from the data which can be read on the chart. The average error thus obtained amounts to 6 per cent in air flow, 7.5 per cent in heat rejection, and 10 per cent in allowable rating. Determining the probable error with Bessel's equation results in slightly smaller errors.

Results of Tests of a 300-Ampere Generator

The prototype of a 300-ampere generator manufactured by Jack and Heintz, Incorporated, type *G-123*, was subjected to the rating tests. The actual test conditions included air-inlet temperatures from −40 to +35 C, pressure altitudes from sea level to 60,000 feet and the pressure drops from 0.7 to 34 inches of water, speeds from 3,500 to 8,000 rpm, and output rating from 50 to 133 per cent. The ratio of pressure drop to outlet pressure, in identical dimensions, ranged from 2.4×10^{-3} to 0.75. Nine different tests were conducted at rated load at the maximum speed of 8,000 rpm, with nine tests at other loads at 8,000 rpm, and four tests at maximum load and lower speeds. Also, a number of duplicate tests were conducted to determine reproducibility of results. The critical component temperatures were 121 C at the drive bearing, as determined by thermocouple measurement, and 205 C in the rotor laminations which were measured by an imbedded thermocouple with the electromotive force transmitted through a specially developed rotating thermocouple switch. Fig. 5, plotting log ($p_a \Delta P$) versus log W, demonstrates that for this machine the temperature effect can be neglected but not the

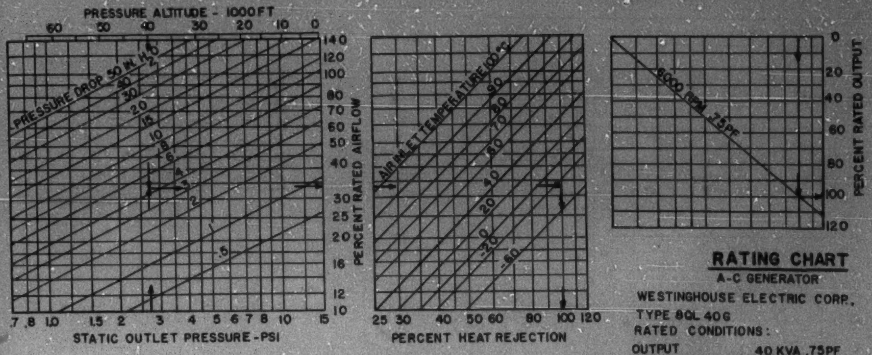


Fig. 4. Rating chart, a-c generator

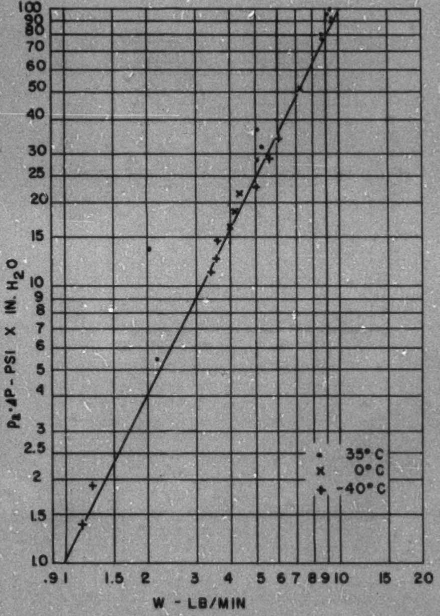


Fig. 5. Air-flow correlation, d-c generator, speed effect neglected

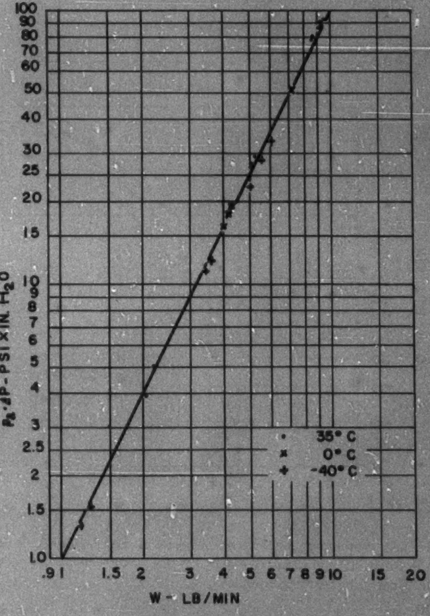


Fig. 6. Air-flow correlation, d-c generator, speed effect incorporated

speed effect, because the extreme deviations from the drawn line refer to conditions with low ratio of pressure drop to outlet pressure. In accordance with the procedure outlined, the speed effect was determined, resulting in the following equation

$$W=\sqrt{p_a\Delta P-\left(\frac{p_a}{4.61}\,\frac{N}{1{,}000}\right)^2} \qquad (18)$$

where

W = air flow, pounds per minute
p_a = outlet pressure, pounds per square inch
ΔP = pressure drop, inches of water
N = speed, revolutions per minute

The final correlation of air flow versus pressure drop and speed is shown in Fig. 6. This relationship was further substantiated by air-flow measurements with a rotameter at different speeds under sea level conditions. For each of the critical temperature limitations, satisfactory heat-

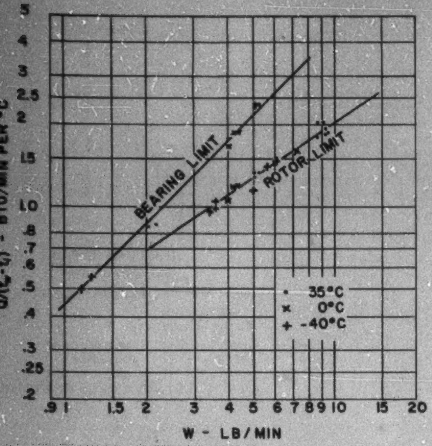


Fig. 7. Heat-transfer correlation, d-c generator

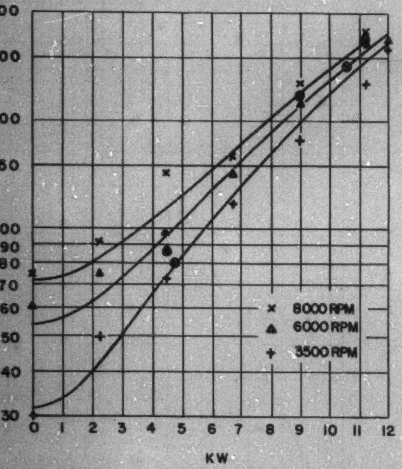


Fig. 8 (above). Loss-load correlation, d-c generator

transfer correlation could be obtained by the following equations representing the bearing and the rotor limitation

Bearing:

$$Q=0.432W(130-t_i) \qquad (19)$$

Rotor:

$$Q=0.432W^{0.\text{[illegible]}}(205-t_i) \qquad (19A)$$

where

Q=heat rejection rate, Btu/min
W=air flow, pounds per minute
t_i=air-inlet temperature, C

The test data and the correlation equa-

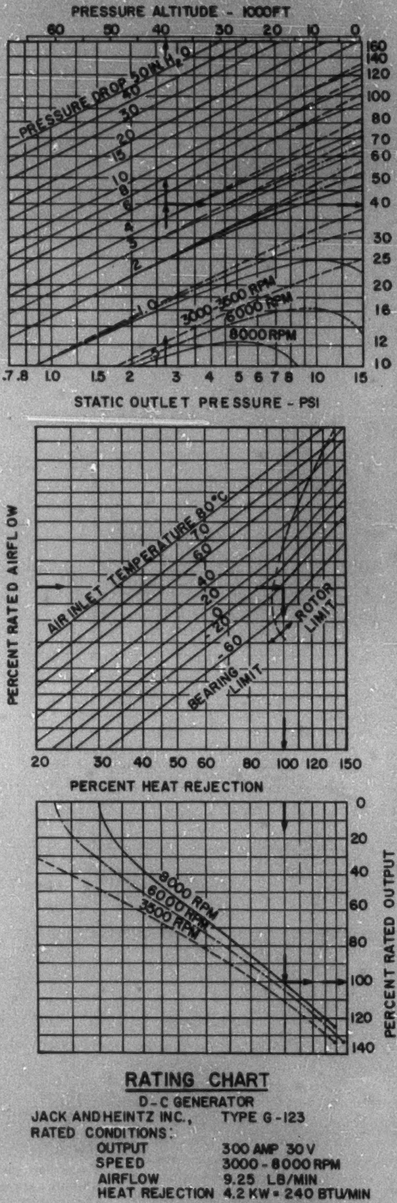


Fig. 9 (right). Rating chart, d-c generator tion are shown in Fig. 7. In the equation covering the rotor temperature limit, the equivalent wall temperature is equal to the temperature limitation of 205 C, whereas in the equation covering the bearing temperature limitation, a value of 130 C was determined because the test results showed the bearing temperature to be approximately 10 C lower than the average outlet temperature. The plot of heat losses versus load is shown in Fig. 8. This plot includes the average values of the data obtained from the rating tests and a number of efficiency tests on a standard drive stand. The rating test data are encircled. In an attempt to formulize the loss data which, especially at low loads, vary considerably because of the uncertain brush friction effect, it was assumed that the losses can be separated into speed and electrical resistance losses by the following equation

$$Q=9\frac{N}{1{,}000}+168\left(\frac{L}{9}\right)^2 \tag{20}$$

where

Q = heat rejection rate, Btu/min
N = speed, revolutions per minute
L = load, kilowatts

The resulting rating chart reduced in scale is shown on Fig. 9. The arrows outline an example of using the charts: at a static pressure of 2.72 pounds per square inch (equivalent to 40,000 feet), pressure drop of 5 inches of water, and air-inlet temperature of −25 C, the per-cent rated air flow is 40 per cent, the per-cent heat rejection is 100 per cent, and the permissible rated load is 100 per cent, which result in an air flow of 3.7 pounds per minute, a heat rejection rate of 240 Btu/min or 4.2 kw, and a permissible load of 300 ampere at 8,000 rpm. An estimate of the accuracy of the chart by means of averaging the absolute deviations of the actual test data results in average errors of $2^1/_2$ per cent in air flow, 4 per cent in heat rejection, and 5 per cent in allowable rating. Using methods similar to those described here, the manufacturer also analyzed the same test data and developed a rating chart essentially in conformance with Fig. 9, but which does not include the thermal limitations of the bearing. Therefore, when using the manufacturer's chart,[16,17] additional calculations must be made which are unnecessary if the bearing limitations are incorporated in the manner shown. Also, the manufacturer's chart incorporates the speed effect on the air flow by a separate plot. Since for this machine the speed effect, although essential, was relatively small, either presentation method is seemingly satisfactory. If the speed effect is of greater influence, the separate plot method will be preferable.

Conclusions

Realizing the immediate need of obtaining adequate rating and application data for aircraft electric equipment and establishing suitable test and correlation procedures, the following conclusions are made:

1. The test and correlation procedure described in this paper establishes the application data of pressure drop, air flow, heat rejection rate, and permissible output for blast-cooled generators within an estimated accuracy of 10 per cent. With more experience and improved instrumentation even closer agreement can be expected.
2. The correlation procedure establishes from a number of tests all essential factors and permits the preparation of a simple graph from which all application data can be taken.
3. The correlation procedure is considered suitable for adoption as a military and industry standard for blast-cooled aircraft generators.

References

1. Cooling Problems of Electrical Equipment in High Speed, High Altitude Aircraft, Erwin Naumann. *Technical Data Digest*, Armed Services Technical Information Agency, Dayton, Ohio, vol. 17, no. 6, June 1952.

2. Study of Aircraft Cooling Systems for Rotating Electric Equipment, Cecil G. Martin. *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 150–60.

3. Equipment Cooling in High Altitude, High Speed Airplanes, R. W. Pfaff. *Document D-10814*, Boeing Airplane Company, Seattle, Wash., Jan. 1951.

4. Cooling of Airborne Electronic Equipment, Vol. I. Ohio State University, Columbus, Ohio, June 1953, "Inadequacies of Equipment Specifications with Regard to Cooling," R. H. Simonds, R. A. Yereance.

5. *Ibid.*, "Installation Cooling Problems of Airborne Rotating Electrical Machines," Leonard J. Lyons.

6. Loss-Temperature-Environmental Relationships for Aircraft Generators, Cecil G. Martin. Jack and Heintz, Inc., Cleveland, Ohio, Aug. 1953.

7. Altitude Rating of Electric Apparatus, Paul Lebenbaum, Jr. *AIEE Transactions (Electrical Engineering)*, vol. 63, Dec. 1944, pp. 955–60.

8. Cooling of Generator Equipment for High Speed, High Altitude Aircraft, E. O. Naumann. *AMC Memorandum Report MCREXE-656-1898*, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, Sept. 1950.

9. Electronic and Electrical Equipment Installation Symposium, Vol. 3. *Bureau of Aeronautics*, Washington, D. C., March 1952, "Cooling of Aircraft Generators," Dale Scott.

10. Altitude Rating of Aircraft A-C Generators, T. F. Hardman, H. Crapo. *Westinghouse Electric Corporation*, Lima, Ohio, Sept. 1953.

11. Introduction to the Transfer of Heat and Mass (book), E. R. G. Eckert. McGraw-Hill Book Company, Inc., New York, N. Y., 1950.

12. Thermal Study of a Blast-Cooled Direct-Current Aircraft Generator, W. Robinson, J. R. Barnum, O. E. Buxton. *WADC Technical Report No. 53-368*, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, Dec. 1953.

13. Rating Tests of a 40 KVA Alternator, E. O. A. Naumann. *WADC Technical Memorandum Report WCLE-53-338, Ibid.*

14. Rating Test of a 300 Ampere Generator, E. O. A. Naumann. *WADC Technical Memorandum Report WCLE-54-20, Ibid.*, Feb. 1954.

15. Devices for Measurement of Rotor Hot-Spot Temperatures of Aircraft Generators by Means of Thermocouples, J. R. Barnum, C. E. Buxton, J. M. Nau, W. Robinson. *WADC Technical Report No. 53-485, Ibid.*

16. Thermal Developments in Aircraft Generators, Cecil G. Martin. Jack and Heintz, Inc., Cleveland, Ohio, Jan. 1954.

17. Generator Developments for High Performance Aircraft, R. J. Eschborn. *Ibid.*, March 1954.

◆

Discussion

Daniel Friedman (Naval Research Laboratory, Washington, D. C.): This paper serves to highlight the valuable contributions which Dr. Naumann has made to the equipment rating field. Of particular merit is the 3-chart presentation method discussed in this paper. There are a few factors, however, which warrant further discussion.

One item of particular importance is the criterion used to evaluate the accuracy of the correlation methods. If one assumes that the wide variation in absolute deviation is caused by experimental inaccuracies, then the average of the absolute deviations is a good measure of the accuracy of the correlation methods. Investigations carried on at the Naval Research Laboratory indicate that this may not be the case. Tests run with instrumentation of known accuracy have shown that large deviations of the same order of magnitude, as listed in Table III, reference 14, have been obtained when using Dr. Naumann's correlation equations, which cannot be attributed to experimental inaccuracies.

Another approach to the problem would be to determine, after subtracting out the known experimental inaccuracies, how accurate the correlation methods are over the entire range of possible operating conditions. In this case, the accuracy of the method is the deviation at the worst conditions.

It is believed that this paper fails to make clear how the mass flow was measured during rating tests. Reference 13 indicates that, for the 40-kva generator, the air flow was measured by an instrumentation section whose calibration was checked only at sea level. Naval Research Laboratory tests of a similar device, when compared to a standard American Society of Mechanical Engineers orifice installation, indicate that there may be an altitude effect on the calibration.

For the 300-ampere generator, reference 14 indicates the generator air flow was calculated from a correlation equation which was based on sea level data only. The

$2^1/_2$-per-cent "error" is, therefore, only the difference between two rating equations and does not reflect the accuracy of equation 18.

To utilize and evaluate experimental rating data, it is necessary to know what environmental test conditions and instrumentation test techniques were used. Dr. Naumann is to be commended for the material of this nature which he has included in his reports. Some of the comments included in this discussion could not have been possible if this had not been so.

Some workers in this field, however, have impaired the value of the data presented by neglecting to provide such information. It is recommended that, until standardized procedures are developed, much more emphasis should be placed on describing in considerable detail how the tests were run.

Erwin O. A. Naumann: As stated in the paper, the data given on accuracy refer to the application data represented by the charts. Thus, the deviations from the actual test data are caused by experimental inaccuracies as well as by the influence of less essential factors either neglected or found to be of minor importance by the correlation method. Since these less essential factors are to a great degree also of a random nature with the greatest influence being the rather unpredictable loss-load relations, it was felt justified to give the data on accuracy in the manner shown, but strictly as an estimate.

In the development of the charts, special care was taken to give the conservative rating for all those cases where deviations in output rating of more than 3 per cent were found. Therefore, the charts give the established safe rating within their readability.

It is correct that ultimately the accuracy of the correlation method is dependent on the deviation at the worst condition. In this connection, it is interesting to note, as shown by Table III of reference 14, and from Tables 1 and 3 of reference 13, that all large deviations in output rating of more than 10 per cent occurred only at extremely low pressure drops across the machine, i.e., 2.0 inches of water or less. Since applications with such low pressure drops are not common, the accuracy appears quite satisfactory and proved to be greater than was expected.

References 13 and 14 describe the methods used in determining the air flow by velocity pressure measurement in a straight 40-inch long, 3-inch-diameter duct section; the measurement being Reynolds-number corrected in accordance with B. Eck.[1] A separate altitude effect was not included since specific data enumerating such an effect are not known to the author. For the 300-ampere generator, the machine itself had to be calibrated because, at the majority of the tests, the duct measurement gave unsatisfactory low readings of less than 1 inch of water, and because prior extensive work at the Naval Research Laboratory indicated this to be a suitable alternate (NRL Reports No. 3961, 4048, and 4355). Mr. Friedman is correct in that the mean deviation of $2^1/_2$ per cent represents mainly the omission of change in temperature; however, these air-flow data were later checked by the manufacturer at their altitude facility using a standard American Society of Mechanical Engineers orifice and reoprtedly were found to be correct. Mr. Friedman's recommendation for more detailed information on experimental rating data are endorsed by the author.

REFERENCES

1. TECHNISCHE STRÖMUNGSLEHRE (book), **B. Eck.** Springer Verlag, Berlin, Germany, pp. 112–13.

2. MEASUREMENT OF AIR FLOW THROUGH AN AIRCRAFT GENERATOR DURING FLIGHT. PART I—INSTALLATION AND CALIBRATION OF INSTRUMENTATION SYSTEM, **Joseph M. Marzolf.** *NRL Report No. 3961*, Naval Research Laboratory, Washington, D. C., April 1952.

3. PART II—ANALYSIS OF FLIGHT TEST DATA FOR F7F-3 AIRCRAFT, **Joseph M. Marzolf, D. Freidman.** *NRL Report No. 4048*, Ibid., Sept. 1952.

4. PART III—ANALYSIS OF FLIGHT TEST DATA FOR F9F-6 AIRCRAFT, **Joseph M. Marzolf.** *NRL Report No. 4355, Ibid.*, April 1954.

Maintenance Testing of Insulation Resistance on Diesel-Electric Locomotives

W. E. KELLEY
MEMBER AIEE

THE Diesel-electric locomotive has been made possible to a large extent by the electric system which so efficiently controls and transmits to the driving wheels and auxiliaries the concentrated horsepower of our modern Diesel engines. This electric system, consisting of generators, motors, switches, relays, meters, resistors, and thousands of feet of wire, depends primarily for its satisfactory operation on the various insulating materials which confine the electric currents to their proper circuits.

The insulating materials on Diesel-electric motive power are subjected to more severe stresses under normal operating conditions than those of almost any other electric system. These stresses, which result in progressive weakening of insulation value, are caused by heat, dirt, moisture, oil, vibration, and aging. Added to these are the cases where the electric systems are stressed beyond their designed limit by unusual operating conditions, such as the necessity of moving a train when one unit of a locomotive has failed.

To determine the weakening of insulation and to prevent as many road failures as possible, a system of periodic checks or tests of insulation resistance of the various pieces of electric apparatus and circuits has been set up. These tests are a part of the regular preventive maintenance schedule and follow much the same pattern for most railroads. The procedure used by one major railroad is given here as a typical example. The schedule of inspection and repairs for road Diesel-electric locomotives is as follows:

1. Daily or turn-around inspection.
2. Semimonthly inspection.
3. Monthly inspection.
4. Quarterly inspection.
5. Semiannual inspection.
6. Annual inspection.
7. Low scheduled periodic maintenance.
8. High scheduled periodic maintenance.

The time between scheduled periodic maintenance shoppings varies with the type of motive power, i.e., passenger or freight, and also with the design of unit. The passenger Diesel-electric locomotive maintenance cycle is alternately one low scheduled maintenance period and one high scheduled maintenance period. Depending on the design of unit, the low period varies from 18 to 24 months, and the high period varies from 36 to 48 months.

The freight Diesel-electric locomotive maintenance cycle is either two or three successive low scheduled maintenance periods depending on the design of unit and followed by one high period. Units covered by the first schedule receive the first low period at 24 months, the second at 48 months, and a high period at 72 months. The other units receive successive low periods at 18, 36, and 54 months, and a high period at 72 months. The schedule for switching-type Diesel-electric locomotives varies somewhat

---

Paper 54-339, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted July 1, 1954; made available for printing July 13, 1954.

W. E. KELLEY is with the Pennsylvania Railroad Company, Philadelphia, Pa.

Acknowledgment is made to L. S. Billau for his co-operation and assistance in preparing this paper.

from that of road locomotives; they receive no semimonthly inspection and scheduled periodic maintenance has been set tentatively at 10-year intervals.

Types of Inspection

DAILY OR TURN-AROUND INSPECTION

A visual inspection is made of all apparatus in electric cabinets, main and auxiliary generators, traction motors, blowers, fans, cab heaters, and defrosters. For Diesel passenger locomotives at a daily or turn-around inspection at the home terminal, a test for power or control grounds is required. For freight and switching locomotives, a test is not required unless trouble is indicated from inspection, the condition of the ground relay, or from operating crew reports.

If a grounded condition of the electric system is indicated, a more thorough check of the insulation resistance to ground is made with a megohmmeter to determine in which part of the power or control circuits the ground is located. In the limited time allotted for a daily inspection, this can be a difficult problem, particularly because of "moisture grounds" in the electric system. These grounds are caused by the large amounts of steam, water, and cleaning fluids which are used for cleaning both the inside and outside of Diesel-electric locomotives and by the extreme sweating conditions of these locomotives during the winter months. Usually, this type of ground disappears when locomotives warm up to operating conditions in service.

In many cases, megohmmeters used as insulation testers are not helpful in differentiating between moisture grounds and solid grounds which will cause trouble. D-c voltmeters have also been used to check voltage drop to ground of control circuits but have not proved successful as they are so sensitive that even a moisture ground will show a large voltage drop. Table I shows the characteristics of several commonly used voltmeters.

Because a quick answer must be given as to whether a Diesel-electric locomotive is available for service, several quick methods for locating grounds are widely used. For control circuits, a 75-volt locomotive cab lamp is used with the circuits energized. Table II shows the characteristics of two types of these lamps. A 50-watt lamp is generally used. Since most moisture grounds cause less than 0.25 ampere leakage to ground, a discernible glow of a test light indicates a serious ground which must be located. For power circuits, which cannot be tested while energized, a bell-ringer set is used with the same results.

Table I. Characteristics of D-C Voltmeters with 75-Volt Scale Used for Checking Control Circuits

	Voltmeter 1	Voltmeter 2	Voltage Values
Internal resistance, ohms per volt	100	200	
Minimum amperes flowing	0.0100	0.0050	75
	0.0066	0.0033	50
	0.0033	0.0017	25
	0.0016	0.0008	12

In general, high-potential tests are not used for trouble-shooting except as a last resort, and then not until motors and generators have been disconnected from the circuit to be tested. When it is deemed necessary, a 1,200-volt high-potential test is made on power circuits after all apparatus has been dried and blown out. In cases where a ground cannot be cleared from a circuit or located in any other way, the "burn" attachment of a high-potential testing machine is used.

SEMIMONTHLY INSPECTION

This inspection is given to passenger and freight locomotives and, as far as electrical testing is concerned, is the same as a daily inspection with the exception that a test must be made for power and control grounds using a megohmmeter. There are more mechanical items of the locomotive to check at this time than during a daily inspection, and during this extra time when the locomotive is out of service a closer check is given to the commutators, brushes, leads, and insulators of all motors and generators, along with a closer inspection and cleaning of the apparatus in electric cabinets.

MONTHLY INSPECTION

This is a basic inspection period for all types of Diesel locomotives, passenger, freight and switching, corresponding to the boiler wash required for steam locomotives. All motors, generators, and electric cabinets are thoroughly inspected and cleaned with an approved cleaning solution, usually petroleum spirits, and blown out with dry air. After this, all power and control grounds must be eliminated, which means that all circuits must show an insulation resistance to ground of at least 1 megohm. A test light, bell-ringer set, or high-potential test machine may also be used to expedite the locating of grounds during this inspection.

Table II. Characteristics of 75-Volt Diesel-Electric Locomotive Cab Lamps

Type	Watts	Amperes at 75 Volts	Minimum Visibility of Lamp: Volts	Minimum Visibility of Lamp: Amperes
25A17/RS	25	0.333	12	0.18
50A19/RS	50	0.666	11	0.24

QUARTERLY INSPECTION

All of the electrical testing performed at previous monthly inspections is required. The principal additional work necessary at this inspection is the checking of the operation and calibration of relays, regulators, thermostats, and alarm and safety switches.

SEMIANNUAL INSPECTION

Again the testing done on previous inspections is repeated. In addition, all exposed wiring is cleaned and painted with an insulating paint and creepage surfaces of rotating apparatus are cleaned and painted.

ANNUAL INSPECTION

All inspection, testing, and repairing performed at the monthly, quarterly, and semiannual inspections are required at this inspection on all passenger, freight, and switching locomotives. At this time, a high-potential test, as required by Interstate Commerce Commission *Rule 253*, must be made on all high-voltage wiring and all high-voltage electric rotating apparatus. This includes all circuits with a potential of more than 150 volts. The test voltage is applied for at least 1 minute between each section of the power circuit and the locomotive frame. For circuits other than generator or motor windings, the test voltage is 75 per cent above normal working voltage; for windings the test voltage is 50 per cent above normal working voltage.

Before this high potential test is given, all electric insulation must be clean and dry and the insulation between any conductor and ground should be measured by means of a megohmmeter which must show an insulation resistance of at least 1 megohm. If the resistance is less than 1 megohm, the faulty circuit must be corrected.

Scheduled Periodic Maintenance

At this time, the locomotive is brought into the back shop where important components are given an overhaul. During the low scheduled maintenance period, many pieces of apparatus are given a heavy repair on the locomotive. For instance, main generators after passing a megohmmeter test have the brush holders removed for check and repair. The commutator is then checked, cleaned, and covered. A thorough visual check is made of the armature and of the fields. If no defects are found, the interior of the generator is sprayed with an insulating paint. Conversely, most apparatus is removed and sent to specialized shops for repair during high scheduled maintenance period.

All circuits are checked with a megohmmeter and must have a resistance to ground of at least 1 megohm. There is an opportunity at this time to check cables, conduit, and wiring which are not normally accessible and any defective wires or cables are replaced or repaired. Dirt which has accumulated in out-of-the-way places is removed and all electric apparatus and cabinets are cleaned with petroleum spirits. A high-potential test as required at annual inspection, is also given.

The repairing and testing done in a modern Diesel locomotive heavy electric repair shop, as outlined by F. Thomas,[1] are much the same today. The policy, generally, is to follow manufacturers' specifications when testing individual pieces of apparatus repaired in this shop. Some deviation from this policy is made in the case of traction motors and generators.

After cleaning, the armatures and stators of traction motors and generators are tested for 3 minutes by motor-driven megohmmeters. If the final reading is not above 100 megohms, the piece of apparatus is given several hours of drying in an oven. If the reading is still low and additional cleaning and drying do not prove successful, repairs, indicated by a high-frequency surge tester or high-potential test machine, are made. When a satisfactory reading above 100 megohms is obtained, the apparatus is given the high-potential test. In rewinding traction motor and generator armatures, many of the high-potential tests after each step of winding have been satisfactorily eliminated by use of a high-frequency surge tester. Many defects are found with the surge tester, which will not show up with a high-potential test. Only those high-potential tests deemed absolutely necessary, including the final test, are made.

Conclusions

1. Additional methods and further development of present methods of testing electric insulation of Diesel-electric locomotives are needed. A portable meter, preferably with a power supply and designed not to indicate grounds below a determined value of leakage current, is needed for maintenance terminals.

2. Standards should be set up for the use of d-c overpotential testing in connection with traction apparatus. This type of testing has been used extensively for large power apparatus but apparently has not been tried by manufacturers or users of traction apparatus. The nondestructive quality of this type of testing is appealing and with proper standards could prove effective for testing the actual circuits of Diesel-electric locomotives as well as for testing individual pieces of apparatus.

3. The electrical maintenance design factors for Diesel locomotives[2] clearly point out the design features necessary for Diesel-electric locomotives so that electric insulating materials will give satisfactory service with a maximum life. For instance, the use of neoprene insulated wire for control circuits caused consistently low megohmmeter readings because of the distributed leakage peculiar to this type of insulation material. Adherence to good design practice will not only increase insulation life but also will aid in the testing of insulation resistance.

References

1. Equipment and Function of a Modern Diesel Locomotive Heavy Electric Repair Shop, F. Thomas. *AIEE Transactions*, vol. 72, pt. II, July 1953, pp. 175–80.

2. Design Factors Favoring Diesel-Locomotive Electrical Maintenance, J. Stair, Jr. *Ibid.*, pp. 155–57.

◆

Discussion

William S. H. Hamilton (Dormitzer Electric & Manufacturing Company, Cambridge, Mass.): I was very much interested in Mr. Kelley's paper in view of the recent development of a device for measuring the resistance to ground from either side of a live d-c circuit, of not over 400 volts potential. Being entirely self-contained with its own power supply, it meets the requirements of conclusion 1 of the paper.

My company has manufactured a ground detector, called the *D-236*, which is calibrated for a range of 250 to 15,000 ohms. The device contains three Mallory *RM-42* mercury cells, which operate a vibrator having a 75-volt 100-cycle a-c output. This goes to the test leads through a 400-volt d-c capacitor which blocks off any d-c flow but permits full passage of alternating current. By means of a suitable resistance arrangement, this alternating voltage is also impressed on a small 2-prong neon glow lamp which goes out quickly when the calibrated dial controlling this resistance is turned to a setting corresponding to the resistance across the test leads.

In operation, it is first necessary to check the calibration. This is done by turning the calibrated dial to the minimum setting and closing the operating and test switches. The latter in effect connects a known resistance across the test leads. The test switch is held closed and the calibrated dial is revolved toward the higher values of calibration until the neon lamp just goes out. A note is then made of the position of the calibrated dial marking corresponding to the known resistance, relative to a vernier scale. This establishes the zero setting for that calibration and the condition of the batteries at that time. All subsequent readings of resistance are then taken with that vernier scale reading as 0 until another calibration is made.

To use the device for testing, one of the test leads is clipped onto ground (the body of the car or locomotive) and the other can be touched or connected to the circuit which is to be tested. The calibrated dial is turned to the minimum position and the operating switch turned on, upon which the neon lamp will light. The calibrated dial is then turned until the neon lamp just goes out, and the resistance to ground is then indicated by noting the graduation on the dial as compared with the 0 setting of the vernier as established in the foregoing.

It should be noted that the neon lamp has a very sharp cutoff, i.e., the light goes out suddenly at the value of the resistance under measurement. The device was then demonstrated at 250, 500, 5,000, and 10,000 ohms resistance respectively.

While this device was developed primarily for use on main line passenger cars, the features described in the foregoing make it very suitable for use on control circuits of Diesel-electric locomotives, which is the reason for mentioning it. The ability to indicate the resistance to ground with the battery still connected to the circuit makes it particularly suitable for a quick check of circuits on daily inspections.

If serious trouble is indicated, clearing up of circuits one by one can proceed until other devices mentioned in Mr. Kelley's paper can be employed.

This *D-236* ground detector is not a highly accurate instrument in the sense that a good ammeter or voltmeter is. It is believed, however, that its accuracy is entirely adequate for the purposes for which it is intended, and that it provides a rugged, fast operating device with practically no parts which can get out of order or break. The batteries are good for about 30 hours of continuous operation and are easily replaceable.

It is intended to reduce the size of this device as soon as some experience has been gained with the present model and the range of resistance to be covered is more fully established. When this is done, it may be possible to put the device into a cylindrical can, somewhat like a flashlight, with a hook to go over the operator's belt.

I would now like to say a few words in defense of solid neoprene insulated wire for voltages of 0 to 150 volts d-c which is rather condemned in conclusion 3 of the paper. I had been in favor of an insulation composed of synthetic rubber with a neoprene sheath which is more or less standard for power cables on Diesel-electric locomotives, and is now the Association of American Railroads (AAR) Specification *581*.[1]

It was found however in the flame tests with the smaller wire sizes no. 14 to no. 8

Table III. Outside Diameters

American Wire Gauge No.	AAR Specification 581: Rubber and Neoprene, Inches	AAR Specification 578: Solid Neoprene, Inches
16	0.210	0.132
14	0.250	0.185
12	0.270	0.206
10	0.290	0.232

American Wire Gauge that the neoprene sheath would split off from the synthetic rubber insulation and that considerable damage would occur to the latter, even though it would not support combustion when the flame was removed. With the solid neoprene insulation, there was no jacket to split off and the damage done by the flame was much less.

As a result, it was decided to adopt this as standard insulation for control wiring on Diesel-electric locomotives for no. 16 to no 8 American Wire Gauge. It was recognized that the insulation resistance would be lower, but it was not believed that this would present any serious problems. This insulation is impervious to water, oil and fuel oil. It is now covered by AAR Specification *578*.[2]

To prevent such wire from being used in high-voltage circuits, Specification *578* calls for the insulation to be red or otherwise color-coded to distinguish it from the synthetic rubber, neoprene-sheathed wire, Specification *581*, which is black in color.

There is another distinct advautage in using this solid neoprene-insulated wire for control circuit work. It is much smaller in diameter than the synthetic rubber with neoprene sheath. This makes it much easier to get into places where many wires have to be crowded together. Table III illustrates this feature for several of the most commonly used sizes.

Incidentally the same flame tests were made at the same time on larger sizes of cable (250,000 circular mils and up) having synthetic rubber insulation with neoprene sheath. These tests failed to disturb the neoprene sheath, probably because of its greater thickness. Finally, as I remember it, an acetylene torch was applied to one piece of cable. Of course, while the torch was in position, there was burning and damage to the insulation but all combustion stopped as soon as the torch was removed. Even this drastic treatment did not cause the neoprene sheath to split.

REFERENCES

1. SPECIFICATION FOR SINGLE CONDUCTOR CLEAN STRIPPING RUBBER OR SYNTHETIC RUBBER INSULATED 0–750 VOLTS D-C NEOPRENE SHEATHED CABLE FOR DIESEL AND ELECTRIC EQUIPMENT. *AAR Specification 581.* Association of American Railroads, Chicago, Ill.

2. SPECIFICATION FOR SINGLE CONDUCTOR NO. 18 AWG TO NO. 8 AWG INCLUSIVE CLEAN STRIPPING NEOPRENE INSULATED 0–150 VOLT CABLE FOR CONTROL WIRING FOR DIESEL LOCOMOTIVES. *AAR Specification 578. Ibid.*

W. E. Kelley: I wish to acknowledge with thanks the interesting and helpful discussion of this paper by Mr. William S. H. Hamilton. The *D-236* ground detector described by Mr. Hamilton is a step in the right direction in providing a quick method to aid electricians at engine terminals in locating grounds on Diesel-electric locomotives.

In reply to Mr. Hamilton's remarks concerning solid neoprene-insulated control wire, it has been found that the distributed leakage of neoprene used as an insulating material has proved very troublesome in Diesel-electric locomotive control circuits. Rubber-insulated, neoprene-jacketed wire has proved much more satisfactory from an insulation standpoint and has not been any more of a fire hazard when used in these control circuits. Most fires on Diesel-electric locomotives would be equally destructive to both types of wire. The smaller size of neoprene-insulated wire is desirable, but there have been no cases where rubber-insulated, neoprene-jacketed wire could not be used because of its size.

Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled Aircraft Generators

D. H. SCOTT
MEMBER AIEE

THE output of electric apparatus is limited usually by maximum allowable temperatures. If electromagnetic and other design factors permit, the output can be increased by applying additional cooling. This increased cooling principle permits greater output per pound of electric apparatus and has been applied to aircraft generators for a number of years to increase electric system capacity and reduce installed weight.

The aircraft velocity provides a pressure source which forces cooling air through the generator via a duct system. Equipment cooled by this forced convection method have become known as blast-cooled units. The maximum power which these generators can deliver continuously depends on the blast-cooling applied, limitations inherent in the generator design, and the environmental conditions which contribute to heat flow by conduction, radiation, and free convection. Integral fans or other auxiliary methods are used to provide cooling during ground operation.

Considerable variation in both blast-cooling and environmental conditions are encountered during aircraft operation. Generators are subjected usually to conditions not described in the generator design specifications. The application engineer desires to provide adequate generator cooling but realizes that excess cooling air imposes unnecessary drag on the aircraft. For these reasons, increasing efforts have been made during the past few years to develop methods of describing the rating characteristics of blast-cooled generators analytically so that measurements taken for one set of conditions can be used to predict the allowable output and other rating factors for other conditions. This paper is an attempt to describe the progress made in accomplishing that objective.

Fundamentals of Rating

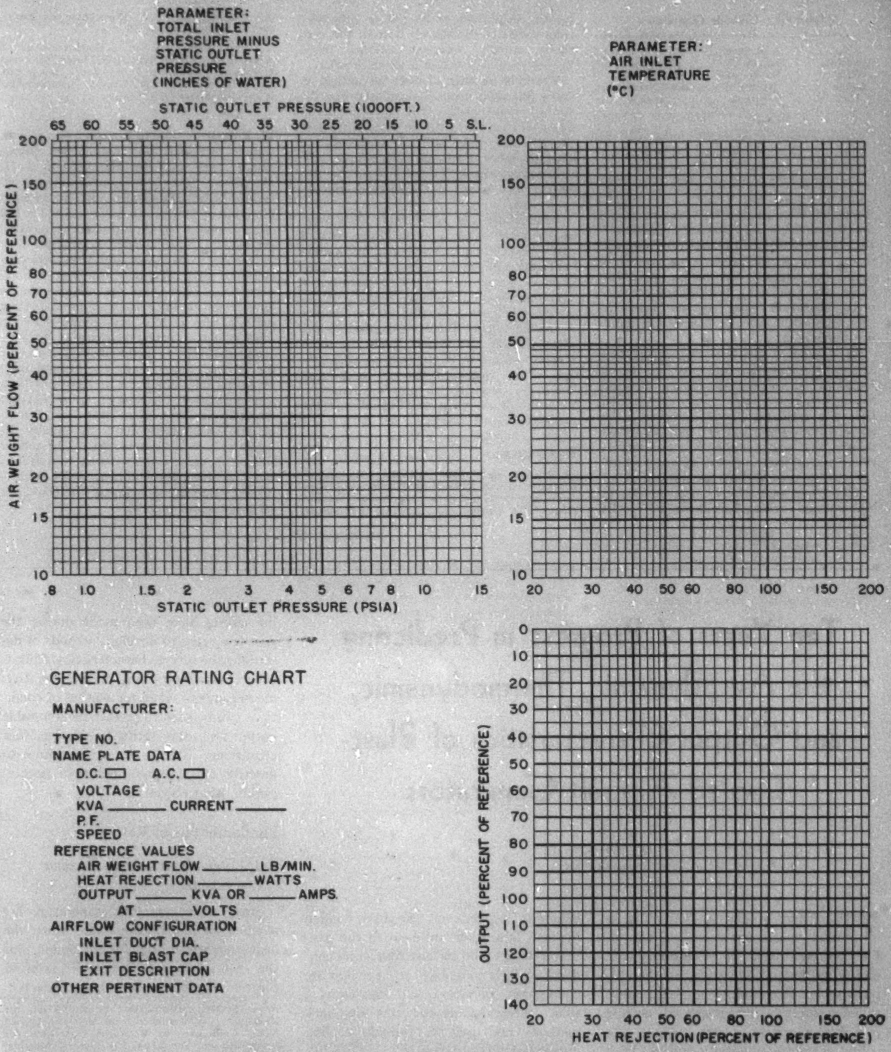
ALL LOSSES REMOVED BY FORCED CONVECTION

Steady-state operating temperatures are reached on electric apparatus when the heat generated within the equipment plus the heat entering from all external sources equals the heat being rejected.

Paper 54-346, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing August 6, 1954.

D. H. SCOTT is with the Naval Research Laboratory, Washington, D. C.

The author, as chairman of the AIEE Subcommittee on Rating of Airborne Electric Apparatus, wishes to acknowledge the contributions of Subcommittee members in establishing the need for this paper. Special acknowledgement is made to the chairman and members of the Blast-Cooling Group of the Subcommittee for many of the cited references. In addition, some of the ideas proposed in the paper resulted from discussion with members of this Group.



GENERATOR RATING CHART

MANUFACTURER:

TYPE NO.
NAME PLATE DATA
D.C. ☐ A.C. ☐
VOLTAGE
KVA ______ CURRENT ______
P.F.
SPEED
REFERENCE VALUES
AIR WEIGHT FLOW ______ LB/MIN.
HEAT REJECTION ______ WATTS
OUTPUT ______ KVA OR ______ AMPS.
AT ______ VOLTS
AIRFLOW CONFIGURATION
INLET DUCT DIA.
INLET BLAST CAP
EXIT DESCRIPTION
OTHER PERTINENT DATA

Fig. 1. Interim data presentation method proposed by AIEE Subcommittee on Rating of Airborne Electric Apparatus, July 1954

In aircraft generator applications, conduction, radiation, free convection, and forced convection contribute to the heat transfer process and their relative contribution depends on the environmental conditions. These environmental conditions, however, are proprietary to the aircraft. It has become common practice, therefore, to assume that environmental conditions are such that all generator losses are removed by forced convection or blast-cooling. Rating characteristics and charts are prepared by the generator manufacturer based on that assumption. The application engineer can then determine the environmental conditions for a particular aircraft and modify the blast-cooling rating as required.

RATING PHASES

The process of rating generators which are cooled by forced convection falls naturally into three phases. The pressure-drop versus air-mass-flow characteristics, the heat-transfer process by which the losses are transferred to the cooling air; and the relation between losses and the allowable generator output. The equations describing these three phases are sometimes combined to form a single rating equation.[1-3] From such equations, generator rating charts have been prepared for use as a reference in applying generators to specific installations.[2-4] An improved 3-chart method has been suggested which presents pressure drop versus air-mass flow, losses, and output data.[5] The primary elements of this presentation method are illustrated by Fig. 1.

Nomenclature

A = area of surface giving up heat to cooling air
F = pressure developed by fan
F_{SL} = pressure developed by fan at density ρ_{SL} and rotor speed N_{SL}
N = rotor speed
N_{SL} = rotor speed at reference sea-level conditions
Q = heat carried away by cooling air
W = mass flow of cooling air
h = coefficient of heat transfer
k = proportionality constant
p_o = static pressure at generator outlet
q_i = inlet dynamic pressure
t_a = temperature of air flowing past surface A (normally considered average of t_i and t_o)
t_i = temperature of inlet air
t_w = temperature of surface past which the cooling air flows
Δp = pressure drop across the generator (total inlet minus static outlet)
Δp_s = static pressure drop across the generator
α = exponent
ρ = mass density of cooling air
ρ_a = average mass density of cooling air (normally considered average of inlet and outlet density)
ρ_i = mass density of cooling air at generator entrance
ρ_{SL} = mass density of cooling air at reference sea-level conditions
σ = density ratio of air (average density divided by average density at sea-level conditions)
σ_1 = density ratio of air at inlet pressure conditions and average air temperature, referred to same values at sea level conditions

Pressure Drop Versus Air-Mass Flow

DYNAMIC PRESSURE AND FIXED RESTRICTION CONCEPTS

For a generator with no integral fan, the following relation between generator pressure drop and inlet dynamic pressure was suggested by C. G. Veinott[6]

$$\Delta p = kq_i \tag{1}$$

It was also indicated that the relation of equation 1 is independent of air density and, hence, aircraft altitude. This dynamic pressure was then related to air-mass flow by

$$W = k\sqrt{\rho_i q_i} \tag{2}$$

This provided one of the earliest means of expressing generator aerodynamic characteristics for a variety of operating conditions from minimum test data.[6]

The well-known orifice or fixed restriction equation for fully established turbulent flow is obtainable by combining equations 1 and 2

$$W = k\sqrt{\rho_i \Delta p} \tag{3}$$

A relation of this form was used as the basis for deriving the equations from which the Bureau of Aeronautics rating chart was prepared.[2]

GENERATOR LOAD CORRECTIONS

The generator losses raise the cooling-air temperature. In addition, the pressure ratio $\Delta p/p_o$ will vary. Martin[7] used a modified form of equation 3 as follows

$$W = k\sqrt{\rho_a \Delta p} \tag{4}$$

which accounts for variation in air density from inlet to outlet.

Barnum, Buxton, and Robinson[8] performed air-flow evaluations on a number of aircraft generators and concluded that, in most cases, flow was found to occur in the region of fully established turbulence. Equation 4, therefore, expressed the flow relation adequately for these generators unless fan effects were present or redistribution of flow occurred. A density ratio was introduced for convenience and σ was substituted in place of ρ_a as follows

$$W = k\sqrt{\sigma \Delta p} \tag{5}$$

E. O. A. Naumann[5] included two methods of correlating air-mass flow versus pressure-drop data on a constant-speed a-c generator. In the first method, changes in air temperature were included, and values of $\sigma_1 \Delta p$ were plotted as a function of W. The equation derived to illustrate the data was of the form of equation 5

$$W \quad k\sqrt{\sigma_1 \Delta p} \tag{6}$$

It was not clear why inlet pressure rather than average pressure was chosen. In the second method, in which air-temperature changes were neglected, values of $p_0 \Delta p$ versus W were plotted. The equation derived to illustrate the data was of the form

$$W = k\sqrt{p_0 \Delta p} \tag{7}$$

Dr. Naumann concluded that it was unnecessary in this case to consider air-temperature changes in determining the air flow. Thus, the relation of equation 7 was used to derive the final air-flow chart.

FAN LAW CORRECTIONS

C. G. Veinott considered the addition of an integral fan and suggested experimentally determined corrections.[6] In a later article[9] it was suggested that the pressure developed by an integral fan could be measured at a reference sea-level condition and then calculated for other conditions by the application of fan laws as follows

$$F = F_{SL}(N/N_{SL})^2(\rho/\rho_{SL}) \tag{8}$$

Combining equations 4 and 8

$$W = k\sqrt{\rho_a(\Delta p \pm F)} \tag{9}$$

This provided a means of including the effects of fan action in the pressure-drop versus air-mass-flow equation.

Barnum, Buxton, and Robinson[8] concluded also that, for some generators, the effect of speed on flow could be expressed as a pressure drop caused by rotation which, when corrected by the blower laws (or fan laws) with respect to speed and flow rate, furnished a single curve of pressure drop versus flow rate.

Marzolf and Friedman[10] plotted values of W^2/ρ_a versus Δp_s for a constant rotor speed. In terms of equation 9, the relation can be expressed as

$$W^2/\rho_a = k(\Delta p_s - F) \tag{10}$$

Thus, by including a pressure correction to compensate for the inherent fan action which opposed the flow, altitude and load effects are describable for this d-c generator.[10] Subsequent data were reported on a different d-c generator and equation 10 describes the relation at all speeds including zero, and at all loads from zero to rated output.[11] The inherent fan effect of d-c aircraft generators has been described by Martin as caused by the open riser type of commutator construction commonly used in the design of these generators which acts like a centrifugal fan and opposes the normal direction of flow.[12] In the application of fan-law corrections, certain rules should be observed. E. O. A. Naumann[5] included an appendix which contained practical suggestions concerning the application of fan laws when these effects are significant.

FLOW REDISTRIBUTION

Barnum, Buxton, and Robinson noted that the relation of equation 9 is not applicable for some generators because of parallel flow paths and flow redistribution with speed.[8] Tests were performed recently at the Naval Research Laboratory on an a-c generator equipped with an integral fan. The application of fan-law corrections did not adequately describe the air-flow changes resulting from rotor speed variation, and a change in flow distribution was detected.

EXCESSIVE PRESSURE RATIO

The incompressible flow equations such as equations 4, 5, and 7 will require corrections for compressibility effects if the pressure ratio $\Delta p/p_o$ becomes sufficiently large. Barnum and Buxton suggested that if $\Delta p/p_o$ exceeds 0.1, a compressibility correction probably will be required.[13] Others have found that this ratio can be considerably higher. Marzolf found that for a $\Delta p/p_o$ of 0.4 no compressibility correction was required for the d-c generator considered.[11]

To explain this discrepancy, it may be noted that most generators consist of several restrictions in series. For many purposes, it is satisfactory to work with the over-all pressure drop Δp but whether or not compressibility corrections are required depends on the pressure ratio across individual restrictions. The correction for compressibility would be required at the lower value of $\Delta p/p_o$ if all of the pressure drop occurred across one restriction.

COEFFICIENT OF HEAT TRANSFER

In forced convection-cooled generators[9] heat is transferred by conduction from its point of origin to the surface past which the cooling air flows and is then transferred to the cooling air. The rate at which heat is transferred between a surface and the cooling air is called the coefficient of heat transfer and is often expressed in watts per square inch per degree centigrade. The situation for steady-state heat flow conditions from a solid having a fixed surface temperature t_w to a fluid flowing past it having a temperature t_a is often expressed as

$$Q=hA(t_w-t_a) \tag{11}$$

The coefficient of heat transfer h depends on units, physical properties of the fluid (air for blast-cooled generators), dimensions of the passages, velocity of the fluid, and other factors.

EFFECT OF ROTOR SPEED

George E. Luke[14] investigated a simulated generator which was cooled by forced convection. The rotor was cylindrical and the air-gap space was uniform. The data indicated that rotor speed seriously affected the coefficient of heat transfer. Paul Lebenbaum[1] discussed the effect of rotor speed on the heat-transfer coefficient and suggested that tests on d-c aircraft generators indicate that the peripheral velocity usually can be neglected.

Barnum, Buxton, and Robinson[8] included values of the coefficient of heat transfer in the various generator passages and for some generators indicated considerable variation in the coefficient with rotor speed. The foregoing results indicate that the effect of speed on the heat-transfer coefficient depends on factors or conditions not defined in the references.

DEPENDENCE ON AIR-MASS FLOW

The generator air-flow path is complex. The nonuniformity of the heat generated and the variables which affect the heat-transfer process suggest the need of an effective heat-transfer coefficient.

Paul Lebenbaum[1] proposed that the effective heat-transfer coefficient depended on air-mass flow alone

$$h=kW^{\alpha} \tag{12}$$

and suggested that α might vary from 0.5 to 1.0 depending on the generator. This is consistent with the relation found to exist for turbulent pipe flow, and for such a relation W. H. McAdams suggested an α of 0.8.[15]

Combining equations 11 and 12 gives

$$Q=kW^{\alpha}(t_w-t_a) \tag{13}$$

Lebenbaum and Veinott both used inlet air temperature in place of average air temperature.[1,9] Equation 13 then is

$$Q=kW^{\alpha}(t_w-t_i) \tag{14}$$

Relations similar to equation 14 have served as the basis for preparing most of the rating charts which have been reported in the literature including the Bureau of Aeronautics chart.[2] E. O. A. Naumann performed a heat-transfer correlation of test data on a constant-speed a-c generator. Values of $Q/(t_w-t_i)$ versus W were plotted on log-log paper and the relation of equation 14 was used to represent the test data.[5]

A similar correlation was performed by C. G. Martin[16] on several d-c generators, and it was indicated that either equation 13 or 14 was applicable but that the use of inlet air temperature, equation 14, gave a slightly better correlation. It was further indicated that no effects on heat transfer were noticeable other than air-mass flow. Since the data were taken over a wide speed range it appears that speed effects on d-c generators may be automatically considered in equation 14.

It should be noted that the data reported by both Naumann and Martin included some deviation from the relation expressible by equation 14. Instrumentation limitations may be a partial explanation for this scatter, but both the complexity of the generator air-flow paths and the complexity of the heat-transfer characteristics are suggested also as possible reasons. Equation 14 is a simplified expression of a complicated phenomenon and can be expected to result in some inaccuracy.

SELECTION OF t_w

Considerable variation has been proposed for the practical selection of t_w. One recommendation is the use of the average of a number of surfaces, another is the use of the surface temperature in the region which limits the rating, while a third is to use the actual hot-spot winding temperature. Insufficient data are available to establish which is preferable.

SHIFT IN HOT SPOT

While the data presented by Naumann and Martin indicate that equation 14 is adequate for expressing the heat-transfer characteristics of some generators, both a-c and d-c, there is still some doubt concerning its general applicability.[5,16] E. O. A. Naumann suggested that the relation of equation 14 may have to be applied to separately distinguishable components of some generators.[5] This would be true, possibly, if the limiting temperature shifted because of a change in condition, such as a change in speed. The limiting temperature might also depend on a definite temperature limit such as a bearing temperature limit.

C. G. Martin indicated that wide variation in the slope of the heat-transfer correlation curves are obtained on a-c generators depending upon the location of the maximum temperature point under consideration. These heat-transfer correlation curves were obtained by plotting $Q/(t_w-t_a)$ versus W on log-log paper, and values of t_w and t_a depended on the temperature point under consideration.[16] Each generator must be evaluated experimentally. If it becomes necessary to apply equation 14 to separately distinguishable components, this may be accomplished by plotting on the heat-transfer chart of Fig. 1 only the relations for the components which are limiting the rating.

ALTITUDE EFFECT

Equation 14 suggests that the heat-transfer characteristics are independent of the outlet static pressure for a given mass flow. Preliminary experimental data obtained at the Naval Research Laboratory on an a-c generator indicate an altitude effect on the heat-transfer characteristic. At least one other laboratory has observed this effect; and both laboratories are conducting additional experiments to determine its magnitude. Additional heat-transfer charts at specific altitudes may be required if this altitude effect proves to be of sufficient importance.

OTHER HEAT-TRANSFER EQUATIONS

Other heat-transfer equations have been developed for describing the heat-transfer characteristics of generators such as the following relation by Crapo and Hardman[8]

$$Q=k\sqrt{\rho_i\Delta p}(t_w-t_i)/[(\rho_i\Delta p)^{0.1}+k_1] \qquad (15)$$

It was recommended that the equation be applied at one speed only.[8] The possibility of fan action suggests that air-mass flow should be substituted for $\sqrt{\rho_i\Delta p}$ as follows

$$Q=kW(t_w-t_i)/(W^{0.2}+k_1) \qquad (16)$$

If k_1 is small with respect to $W^{0.2}$, equation 16 becomes

$$Q=kW^{0.8}(t_w-t_i) \qquad (17)$$

which is of the same form as equation 14.

Loss Versus Output

There is no single fixed relationship between output and losses which holds for all aircraft generators. It has been customary to divide the losses into two broad categories; constant losses and variable losses.

D-C GENERATORS

For d-c generators, the variable losses vary usually as the square of the current for a given speed. C. G. Veinott[9] suggested that above rated load the total losses varied approximately as the square of the load current, while below rated load it was more correct to consider that the losses vary as the first power of the load current.

Rating equations used to prepare combined rating charts were simplified usually to require the total losses to be proportional to some power of the load current. Paul Lebenbaum[1] assumed the losses to vary as the square of the load current to derive his simplified rating equation and justified it on the basis of the relative magnitude of the armature circuit losses as compared to the total. To derive the Bureau of Aeronautics rating chart, it was assumed that the losses varied as the 1.5 power of the load current.[3] This combining of the fixed and variable losses to simplify the equations limited the application of the rating charts to outputs above approximately 40 per cent.

A-C GENERATORS

In considering a-c generators, the power factor is important and influences the loss-output relation. It may be sufficient in some cases to consider only rated power factor conditions. The complex air-flow paths and the use of integral fans on a-c generators are probably the cause for redistribution of air flow for various conditions. If it becomes necessary to apply equation 14 to separately distinguishable components of a generator, as was suggested in the section entitled "Heat-Transfer Progress," the loss versus output curve can be plotted for the conditions which are limiting.

SPEED EFFECTS

It is recognized that generator losses vary with speed. The data presentation method illustrated by Fig. 1 permits the use of speed as a parameter in the loss-versus-output chart and provides more accurate information than was permissible on the combined-type rating charts. If loss-versus-output data are required for variation in both speed and power factor, additional output charts will be required.

VARIATION IN LOSSES WITH ALTITUDE

It has been customary to evaluate generator losses at sea-level pressure conditions and apply these same values at altitude conditions. This approach has been justified by the argument that losses at sea level equal or exceed the losses at altitude. Preliminary test data taken at the Naval Research Laboratory on an a-c generator indicate the losses at 50,000-feet pressure altitude are approximately 20 per cent less than the sea-level losses. This loss variation is sufficient to justify measurement of the losses at altitude as well as at sea-level conditions in order to establish the relation for specific generators.

Minimum Test Data

The minimum data required to insure that the three phases of rating are described adequately will vary with the generator and the possible application. As indicated in a previous article, the information required to derive the Bureau of Aeronautics rating charts was more in the form of minimum specification values.[2]

EXPERIMENTAL CORRELATIONS

The mass-flow, heat-transfer, and output correlations included in E. O. A. Naumann's report[5] contained approximately 30 data points, and it appeared that sufficient instrumentation was installed so that data for all three phases were obtained simultaneously. It is to be noted that this was a constant-speed generator. Many additional tests would have been required to evaluate the effect of speed.

The heat-transfer correlations included in C. G. Martin's paper[16] contained from 20 to 40 data points for each d-c generator. As indicated earlier, Martin suggested that no effects on the heat-transfer coefficient were noticeable other than the air-weight flow on these d-c machines. It would be expected that additional tests were required on the air-weight-flow correlation to evaluate the speed effect suggested, but the data points were not indicated on the air-weight-flow correlation.

The heat-transfer correlations included in Barnum and Buxton's paper[13] were made for the stator, rotor, and air-gap surfaces separately. Each of these correlations included from 12 to 15 data points. An over-all heat-transfer correlation was not indicated.

BASIC CONSIDERATIONS

The determination of minimum data points should include consideration of the following items: 1. Data should be taken at definite design points which are of specific interest; 2. it is customary to require at least three points on a curve to establish a trend; 3. data points should be selected so that interpolation rather than extrapolation prevails.

Consistent with the foregoing factors, the minimum data points can be readily determined once the variables have been established. As additional test results become available from the various organizations which are now conducting concentrated rating studies, the relative importance of speed and altitude effects will be clarified and the minimum test data required to determine rating adequately can be established.

Instrumentation Developments

ROTOR TEMPERATURE MEASUREMENT

The rotor is often the temperature-limiting component on aircraft generators. Temperatures taken after shutdown have little meaning because of the resulting temperature redistribution. In addition, the variety of conditions encountered in generator operation require temperature measurements during operation. The study reported by R. F. Sims[17] indicated that thermocouples furnish the most satisfactory method for measuring generator rotor temperatures during operation. Various attempts have been made to attach extra slip rings and brushes to generator rotors for this purpose, but the compactness of aircraft generators has proved a severe limitation.

Considerable progress has been made on the development of devices mounted on separate bearings which can be connected to revolving thermocouples and mechanically driven by the generator rotor. P. R. Tarr[18] described a device developed to measure the temperature of rotating turbine blades. C. G. Martin described a similar device which has been developed for aircraft generator application.[12] More recently a device using mercury coupled elements has been developed.[19] The data resulting from generator thermal evaluation using these new types of devices for measuring rotor temperatures are serving as a basis for establishing experimental techniques in order to determine generator rating charts.

Conclusions and Recommendations

For a given speed the air-mass-flow versus pressure-drop characteristics of most generators can be described by equations applicable to air flow through a fixed restriction. The effect of generator load can be considered by using an average or effective air density; see equation 4. In some cases the changes in air temperature can be neglected and the outlet static pressure used in place of average density; see equation 7. If pressure sources are present because of an integral fan or inherent fan effect, corrections are required in the equations to describe the flow correctly. On some generators fan-law corrections can be applied; see equation 9. When redistribution of flow is significant, the fan-law corrections may not be adequate.

Since the air-flow path consists of several restrictions in series, the pressure drop across any one restriction is a fraction of the total, and the use of incompressible flow theory can be used for larger values of $\Delta p/p_o$. For some generators the effective heat-transfer coefficient depends only on air-mass flow, as in equation 12. For these generators the following relation can be used to illustrate the data

$$Q = kW^{\alpha}(t_w - t_1) \qquad (14)$$

Considerable variation prevails concerning the choice of t_w. The hot spot which limits generator output shifts in some generators because of the complex air-flow path and the nonuniformity of heat generation. In such cases equation 14 may have to be applied to separate components rather than to the complete generator.

An altitude effect has been observed in the heat-transfer characteristics of some a-c generators. This requires further investigation. The proposed data presentation method illustrated by Fig. 1 permits the use of either power factor or speed as a parameter in the loss-versus-output chart. Additional charts can be added if both parameters are to be varied. If the hot spot shifts with conditions, the loss-versus-output curve can be plotted for the conditions which are limiting. The minimum data points required can be readily determined once the variables are established. The development of instrumentation to measure rotor temperatures during operating conditions provides a means of obtaining realistic rating data.

Equations 4, 7, 9, and 14 coupled with Fig. 1 represent major accomplishments in the attempt to describe the thermal rating characteristics of blast-cooled generators analytically. These equations are simplified expressions of complex phenomena but can be used by the application engineer as a reference.

Experimental data should be obtained on each generator design to determine analytical expressions which can be used to describe approximately the aerodynamic, thermodynamic, and loss-versus-output characteristics. If practicable, these data should be taken in regions which will permit the use of analytical expressions to interpolate between data points rather than extrapolate beyond data points.

References

1. Altitude Rating of Electric Apparatus, **Paul Lebenbaum, Jr.** *AIEE Transactions (Electrical Engineering)*, vol. 63, Dec. 1944, pp. 955–60.

2. Cooling of Aircraft Generators, **Dale Scott.** *Electronic and Electrical Equipment Installation Symposium, Bureau of Aeronautics*, Washington, D. C., vol. 3, March 1952, pp. 25–34.

3. Altitude Rating of Aircraft A-C Generators, **Hayes Crapo, T. F. Hardman.** *AIEE Special Publication S-57*, Oct. 1953.

4. Blast-Cooled Aircraft Generator Rating Curve. *Bureau of Aeronautics*, Washington, D. C., drawing no. 49A1B15, revised May 1952.

5. Rating Tests of a 40-KVA Alternator, **E. O. A. Naumann.** *WADC Technical Memorandum Report WCLE 53-338*, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, Dec. 1953.

6. Blast-Tube Cooling for Aircraft Generators, **Cyril G. Veinott.** *AIEE Transactions (Electrical Engineering)*, vol. 63, July 1944, pp. 520–25.

7. Study of Aircraft Cooling Systems for Rotating Electric Equipment, **Cecil G. Martin.** *AIEE Transactions*, vol. 71, pt. II, July 1952, pp. 150–60.

8. Heat Transfer Processes within Rotating Electrical Equipment for Aircraft, **J. R. Barnum, O. E. Buxton, W. Robinson.** *R. F. Project* 445, Ohio State University Research Foundation, Columbus, Ohio, report no. 3, July 1952.

9. Effect of Altitude on Ventilation and Rating of Aircraft Electric Machines, **Cyril G. Veinott.** *AIEE Transactions (Electrical Engineering)*, vol. 65, Feb. 1946, pp. 84–90.

10. Measurement of Air Flow Through an Aircraft Generator During Flight—Part II, **J. M. Marzolf, D. Friedman.** *NRL Report 4048*, Naval Research Laboratory, Washington, D. C., Sept. 1952.

11. Measurement of Air Flow Through an Aircraft Generator During Flight—Part III, **Joseph M. Marzolf.** *NRL Report 4355*, Naval Research Laboratory, Washington, D. C., April 1954.

12. Loss-Temperature-Environment Relationships for Aircraft Generators, **Cecil G. Martin.** *AIEE Special Publication S-57*, Oct. 1953.

13. Temperature Distributions and Heat Transfer in a Blast-Cooled Generator, **J. R. Barnum, O. E. Buxton.** *Paper No. E-6*, Ohio State University, Columbus Ohio, "Cooling of Airborne Electronic Equipment," vol. 1.

14. The Cooling of Electric Machines, **George E. Luke.** *AIEE Transactions*, vol. 42, June 1923, pp. 636–52.

15. Heat Transmission (book), **William H. McAdams.** McGraw-Hill Book Company, Inc., New York, N. Y., 1942, p. 166.

16. Thermal Developments in Aircraft Generators, **Cecil G. Martin.** Jack and Heintz, Inc., Cleveland, Ohio, Jan. 1954.

17. The Measurement of the Temperature of a Rotating Armature, **R. F. Sims.** *Technical Note EL-11*, Royal Aircraft Establishment, Farnborough, England, May 1950.

18. Methods for Connection to Revolving Thermocouples, **Philip R. Tarr.** *NACA Research Memorandum RME50J23a*, National Advisory Committee for Aeronautics, Washington, D. C., Jan. 1951.

19. Devices for Measurement of Rotor Hot Spot Temperatures of Aircraft Generators by Means of Thermocouples, **J. R. Barnum, O. E. Buxton, J. M. Nau, W. Robinson.** *WADC Technical Report 53-485*, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, Nov. 1953.

◆

Discussion

Robert A. Yereance (Boeing Airplane Company, Seattle, Wash.): I have tried to plot data on an alternator tested on the curves of Fig. 1. I feel that the curves present effectively a large amount of data and so are definitely useful. However, the data indicate that curves of air weight flow versus heat rejection must be drawn for each altitude. This does not necessarily decrease the usefulness of the curves but it should be recognized that data taken at a particular altitude condition cannot be used for all operating conditions in every case. Therefore, for each type of machine, check points should be taken over a range of altitudes to determine if one set of sea-level curves will suffice or if complete altitude data must be taken. This paper does an excellent job of summarizing the history of the problems of blast-cooled aircraft generators and thus makes very good reference material.

D. H. Scott: The comments submitted by Mr. Yereance are consistent with the findings discussed under the subheading "Altitude Effect." Of particular interest is that this altitude effect on the heat-transfer process has been observed on at least two different manufacturers' generators. I concur with Mr. Yereance that, for each type of machine, check points should be taken over a range of altitude to determine if one set of sea-level curves will suffice or if complete altitude data must be taken.

Operating Experience in Diesel-Electric Locomotives Results in Design Changes

L. ELTON LEGG
MEMBER AIEE

THERE is no question regarding the soundness of engineering employed during the design and construction of Diesel-electric locomotives, but many of the refinements follow as a result of the experiences gained by the various railroads operating them. This power plant on wheels was intended to outperform a steam locomotive and it will, but there was a day when it would not. The iron horse was little affected by the weather but many of the earlier Diesel-electric locomotives could not be operated successfully in such minor weather adversities as mist or fine snow.

Ground Relay Action Due to Nuisance Grounds

The ground relay is a very important piece of equipment on a Diesel-electric locomotive but experience gained by the railroads has proved that it must not be oversensitive. Some ground relays were first designed to trip on such low current leakage that grounds in the high-voltage system, created by mist or fine shifting snow, would render the locomotive inoperative. This plagued the railroads to such an extent that the term "nuisance ground" was coined. Such a term was indeed fitting, as the writer has seen as many as nine passenger locomotives out of service at one time as a result of such grounds. After railroads desensitized their own ground relays with satisfying results, some of the builders standardized on a tripping value as high as five times the original. No ill effects were observed because of the desensitization. Another very important step in the right direction was in the adoption of neoprene-insulated wire and cable. The original varnished cambric type of insulation permitted moisture creepage which caused nuisance grounds. This change was made because of the many complaints made by the railroads.

Moisture in Main Generator Damages Windings

Another great source of trouble in the Diesel-electric locomotive was grounded interpoles in the lower area of the main generator which also caused ground relay action. Because of the cooling air being taken from without the car body, it was often laden with fine snow and moisture which was driven into the main generator. Such moisture collected carbon dust, given off by the brushes, and finally found its way down around the three lower interpoles and caused premature decomposition of the insulation. After reaching a certain state, current leakage through the weakened insulation and carbon deposits would begin causing tripping of the ground relay, and often the main generator would have to be replaced.

Annual Dielectric Test Damages Windings

One of the greatest maintenance costs was found in main generator and traction motor replacements resulting from the annual dielectric test required by the Interstate Commerce Commission.[1] This yearly test of the insulation has to be made in a prescribed manner and the railways, of course, carry it out faithfully. Even though the greatest care is employed in performing the dielectric test, many times the insulation is damaged by such test to require removal of traction motors or the main generator. It is not such a sizable task to interchange traction motors but such is not the case with main generators. In fact, to change a main generator, several heavy pieces of auxiliary equipment must first be removed and the roof of the locomotive lifted. The complete change costs $3,000 $4,000 $5,000 or $6,000 in addition to losing the use of a $200,000 locomotive for not less than 3 days.

Dielectric testers as known in the past were designed mainly for testing new or very good insulation. A common use for these testers was to cull out defective portions of electric apparatus to prevent their breaking down after reaching the hands of the purchaser or user, thereby protecting the reputation of a manufacturer. This culling-out process was often done at the expense of replacing insulation but it has always been considered common and good practice. The necessity or importance of such a test has not changed with the arrival of the Diesel-electric locomotive as the practice is being carried out by all locomotive builders. It is also equally important in railway armature rewind shops; however, after traction motors and generators have once been installed in the locomotive, the point of view definitely changes. The one and all-important issue is, then, to give such equipment the proper care to insure long and dependable life. The result of using dielectric testers, which were designed primarily for testing new insulation, has brought about the absolute necessity of newer and better methods less destructive to the electric equipment. Certainly, the Interstate Commerce Commission did not intend to impose such costly maintenance upon the American Railroads when rule no. 253 was written; otherwise, the last sentence of this rule would never have been included. This last sentence reads as follows and pertains to after test: "A careful examination shall be made of any weakness indicated and all defects remedied before locomotive is put in use." Most methods of dielectric testing do not give any indication of weakness, but, rather, further damage the insulation to a point to which it could not possibly be remedied except by removal.

Paper 54-344, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing July 16, 1954.

L. ELTON LEGG is with the Chicago & North Western Railway, Chicago, Ill.

Here, again, it was usually the insulation on one of the three lower interpoles which would break down under the stress of the required test. This problem was confronted from two different angles: the first, to provide immediate relief, and the second, to provide relief from a long-range standpoint.

Long-Range Relief

The long-range program was to provide a barrier of hard insulation between the interpoles and frame of the main generator as well as to improve the insulation on the winding itself. Here, again, the builders have followed suit. Some railroads have altered the ventilating system so as to take the cooling air from within the car body, thereby eliminating the moisture. Correction has also been effected by the builder in the latest type of locomotives. The main generator being nearly a solid mass of metal is a very difficult machine to clean. If an air stream is played into the open end, the carbon dust and dirt will move about somewhat but only a small portion of it will actually find its way out. Through necessity, some railroads have drilled holes in the generator fan which permits the blowing of air from the engine end of the generator driving the dirt and carbon dust outward. This has proved to be a very effective means and it is expected that the builders will take some action in this respect. To insulate the interpoles, alter the ventilating system, and drill the generator fans on a fleet of Diesel-electric locomotives, several years would be required as such work could only be accomplished at disassembly time; consequently, steps for immediate relief were necessary.

Immediate Relief

Immediate relief was brought about by designing a nondestructive dielectric tester solely for railroad use. Rule no. 253 specifies the potential requirements as well as the duration of the annual test. Even though most dielectric testing is performed with alternating current, the Interstate Commerce Commission does not specify the nature of the current to be used; therefore, it was decided to take advantage of the benefits afforded through the use of direct current.

The following requirements were considered pertinent in the planning and development of a dielectric tester which is thought to be most practical for railroad use:

1. It should be capable of impressing the dielectric stress upon the insulation as required by law and, in addition, should actually anticipate an insulation breakdown before it occurs, thus permitting the railroads to take advantage of the provision in the last sentence of rule no. 253.

2. It should be designed so that the operator is not required to hold the test prods in the hands, as the required potential is considered dangerous, especially where perspiration is present. Further, once the high-voltage circuit is made, it should not be interrupted until after the test has been completed.

3. An electric instrument should register actual voltage impressed upon the insulation.

4. The energy for its operation should be taken from the locomotive battery, thus permitting its use, irrespective of location and availability of a-c power.

5. The division between good and rejected insulation should be definite and clearly indicated by an electric instrument and not left to the judgment of an individual.

6. The value or intensity of the leakage current at the division point between good and rejected insulation should be comparable to that of dielectric testers commonly accepted by the trade.

7. The device should take its energy from a 32, 64 or 128-volt battery which is the potential of batteries commonly used on Diesel-electric locomotives and should automatically select the proper voltage for its use.

8. It should be equipped with a 60-second timing device to indicate the duration of the test.

The resultant dielectric tester nearly meets the foregoing requirements. The actual impressed direct voltage is indicated by a voltmeter on the panel of the tester. The intensity of the current leakage through the insulation is indicated by a second instrument whose dial is divided into two clearly defined areas, one colored green and marked "good," and one colored red and marked "reject." If the hand continues to swing upward as the operator gradually increases the voltage, he knows the insulation is defective and immediately reverses the rheostat to reduce the current flow before damage is incurred. After the defect has been located, that particular area may often be repaired. If the insulation is serviceable, the hand will swing upward only a short distance and stop. The operator then continues to increase the voltage to the test requirements. When this value is obtained, he switches a timer on and after 1 minute has elapsed a bell rings and the test is completed. This nondestructive tester does not further damage decomposed insulation. It simply indicates thats uch conditions are present.

Flashovers

Another very costly and troublesome item on Diesel-electric locomotives is the damage resulting from main generator flashovers. Much discussion on this subject has taken place but it seems that the definite cause has not yet been pinpointed. Flashovers occur under any and all conditions and at times the intensity is of such great violence that the armature is damaged beyond further use. Invariably porcelains are shattered and, in some cases, pieces of porcelain have lodged between the brush holders and commutator causing great damage to the commutator. Usually, the brush holders must be removed for repairing. On some of the newer locomotives, the porcelains have been eliminated and replaced by a plastic insulation which does not seem to shatter under stress. Such plastic insulation over the studs does, of course, end the shattering of porcelain but the flashovers continue. The most common path followed by flashover current is from the heel of a brush holder to the frame of the generator and back to the heel of a brush holder of opposite polarity. During this burst of energy, it is not uncommon that the damage requires 2 or 3 days to repair.

One recent railroad development being tested by more than 15 railroads is a neoprene cover moulded to fit snugly over the heel of the brush holder. This provides a barrier between the heel of the brush holder and the frame of the machine. So successful has the test been that 100 locomotives are being completely equipped by one railroad and a substantial number by several others.

Fuses—Source of Trouble

Fuse holders built into the doors of high-voltage cabinets have been a source of great trouble, as fuses falling out of the holders and down into the high-voltage circuits frequently cause serious delays. Another difficulty is that enginemen will not dispose of burned-out fuses but, instead, place them back into the holders, often causing delays when spare fuses are needed. One means for correcting both conditions is for the builders to provide a combination fuse holder and tester with the various size fuse clips connected in series with a push button, indicating lamp, and a source of battery energy. This would permit the enginemen and maintenance men to test all spare fuses in an instant by simply pushing a button; it is even better to use circuit breakers throughout.

Interlocks Improved

Interlocks have probably caused more road trouble than any other single item on a locomotive because they constitute the open-type contact which collects dust, causing poor electric contact. Here, again, experience gained through operation has revealed the necessity of improved interlocks and the builders have again brought relief with the inclosed type of interlock; however, thousands of the older type of interlock are still being used. One fast and effective means for checking interlocks for conductivity is through the use of a low-reading ohm ohmmeter which, of course, should read zero if the interlock being tested is in proper condition.

Conclusion

The foregoing items are only a few of the many Diesel-electric difficulties, some of which have been overcome through close co-operation between the railroad and the builder. The remarkable progress made by the various locomotive builders over the past decade is undoubtedly due to the competitive system of enterprise in combination with the railroad people's informative assistance. There is no greater necessity in railroading than keeping the locomotives in operation.

Reference

1. Laws, Rules and Instructions for the Inspection and Testing of Locomotives Other than Steam. *Rule No. 253*, Interstate Commerce Commission, Washington, D. C.

Discussion

H. F. Brown (Gibbs & Hill, Inc., New York, N.Y.): In the second paragraph the author states, "Another very important step in the right direction was in the adoption of neoprene-insulated wire and cable" for the elimination of nuisance grounds caused by moisture creepage on varnished cambric type of insulation. In a paper by W. E. Kelley,[1] it is stated that "the use of neoprene insulated wire for control circuits caused consistently low megohmmeter readings because of the distributed leakage peculiar to this type of insulation material." These two statements seem to be at variance. Would the author care to amplify his views on this matter?

REFERENCE

1. Maintenance Testing of Insulation Resistance on Diesel-Electric Locomotives, W. E. Kelley. *AIEE Transactions*, vol. 73, pt. II, 1954 (Jan. 1955 section), pp. 452-55.

L. Elton Legg: My reference to neoprene-insulated wire and cable was very vague. I actually had in mind the type of neoprene insulation which is applied by the strip process and which insulation has natural up-river fine para rubber next to the conductor. My experience with wiring on Diesel-electric locomotives has been mainly with this type of insulation which has given us no trouble of the nature referred to by Mr. Kelley. No doubt Mr. Kelley's statement is correct; however, I am unfamiliar with low megohmmeter readings resulting from neoprene-insulated wire.

A Weight Analysis of Modern Aircraft Electric Systems

H. L. GARBARINO
ASSOCIATE MEMBER AIEE

A. K. HAWKES
ASSOCIATE MEMBER AIEE

J. A. GRANATH
ASSOCIATE MEMBER AIEE

THE problem of choosing suitable electric systems for aircraft has been the subject of investigation and discussion for many years. To meet general or specialized requirements, the aircraft industry has had to choose between primary a-c and d-c systems. The magnitude of system voltage has been a variable. For the a-c system, a choice must be made between variable and constant frequency, and for the latter the selection of a nominal value and tolerances are required.

The purpose of this paper is to present recent weight comparisons and to discuss other considerations which are important in the selection of an electrical system. Because of the constant change in aircraft requirements and, consequently, the change in philosophy of design, it is most desirable that the electric system be continually re-evaluated. In this way, the power requirements of an aircraft may be met in a satisfactory manner with minimum delay between the recognized need and the accomplished change.

The data given in this paper are part of the results of an extensive analysis and comparison of electric systems. The investigation was principally directed toward the evaluation of higher voltage d-c systems but, although the emphasis was placed on the one system, its merits and faults could not be gauged except by comparison with others. The three systems studied were the 120-volt d-c, the 30-volt d-c, and the 120/208-volt 3-phase 400-cycle a-c systems.

It is believed that the most important criterion for a comparison of systems is weight, and that lesser criteria are reliability, ease of operation, personnel safety, simplicity, and maintenance. This viewpoint is offered because weight is the factor which can most readily be expressed in concrete terms and because any system can somehow be made satisfactory from the standpoint of the other factors. It is appreciated that some types of aircraft systems present difficult engineering problems for achieving satisfactory operation. However, past experience indicates that a pronounced weight advantage provides an incentive to overcome technical difficulties and to make a new system comparable in terms of factors other than weight.

Weight Comparisons of Systems

The cable weight, to transmit a given amount of power, is identical for the higher voltage d-c system and the 120/208-volt d-c system except for the slight difference caused by the power factor of the a-c system. This tends to make the a-c system heavier than its equivalent higher voltage d-c system. However, the average steady-state power factor of such an a-c system is about 0.95, so the cable weight difference between the higher voltage d-c system and the a-c system caused by power factor is neglected. Cable weight for the 30-volt d-c system is, of course, the highest of the three.

Paper 54-357, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 21, 1954; made available for printing August 5, 1954.

H. L. Garbarino, A. K. Hawkes, and J. A. Granath are with the Armour Research Foundation of Illinois Institute of Technology, Chicago, Ill.

This paper is part of a study carried out for the Electrical Branch of the Equipment Laboratory, Wright Air Development Center, under Contract No. *AF 33(616)-142*.

The a-c system requires a constant-speed drive which adds considerable weight to this system. However, the generator weight of the a-c system is less than that for a d-c system of similar capacity, since it operates at a speed of 6,000 rpm, while the d-c machine must be designed for speeds as low as 3,000 rpm. As an example, to show this weight comparison, consider a 30-kw system. A 40-kva 30-kw alternator weighs approximately 80 pounds, while the constant-speed drive for use with this machine weighs about 90 pounds. A 30-kw d-c generator weighs approximately 140 pounds. Thus, the drive-alternator combination weighs only about 30 pounds more than the equivalent d-c generator.

There is also another consideration, often overlooked, which should be included in the system weight comparison. This is the weight of fuel required to supply the system losses. A method for computing these fuel weights from system efficiencies is presented.

Fuel Weight Comparisons

The weight of fuel used to produce useful work in the electric load equipment can be traced to the function performed by the load device. For given functions to be performed on an airplane, this fuel weight is independent of the type of accessory system supplying power to the devices. Since the amount of fuel used to perform the load functions is the same for each type of electric system to be considered, this part of the fuel weight can be omitted from the comparison. Where the systems compared have different efficiencies, however, the fuel weight caused by losses is related to the type of systems being considered and should be included in the comparison.

Table I. Comparison of Losses for the Three Electric Systems

Item	Approximate Efficiency, e	Losses $\left(\frac{1}{e}-1\right)\times$ Rating, Per Unit
30-volt d-c system		
Generator	0.9	$(1.11-1.0)(1.0)=0.11$
Cable (reference)	1.0	
Inverters	0.6	$(1.67-1.0)(0.1)=0.067$
Total		0.177
120-volt d-c system		
Generator	0.9	$(1.11-1.0)(1.0)=0.11$
Cable (reference)	1.0	
Inverters	0.6	$(1.67-1.0)(0.1)=0.067$
Total		0.177
120/208-volt 3-phase a-c system		
Constant-speed drive	0.9	$(1.11-1.0)(1.11)=0.12$
Alternator	0.9	$(1.11-1.0)(1.0)=0.11$
Cable (relative to d-c)	0.95 (due to power factor)	$(1.05-1.0)(1.0)=0.05$
Transformer-rectifiers	0.95	$(1.05-1.0)(1.0)=0.005$
Total		0.285

The losses to be accounted for are those of the cables and generators of all systems, the constant-speed drive and transformer-rectifier conversion equipment of an a-c system, and the rotary inverters of a d-c system. Table I indicates approximate orders of magnitude of losses of these items in terms of system ratings.

When a figure for the losses is available, the next step is to convert the power figure to a weight of fuel required to supply these losses. A factor which indicates the energy delivered by a prime mover in terms of pounds of fuel is the specific fuel consumption, or marginal fuel consumption, in pounds per kilowatt-hour. The specific fuel consumption is defined as the average of fuel consumption of an engine for a particular type of fuel over the full range of engine operating conditions. A more realistic figure is the marginal fuel consumption, which is the increment of fuel required to increase the engine output by an increment of power for supplying the electric system. This is merely the increment of power above that required to maintain flight in a given flight condition. Thus, the nonlinear fuel-to-output relationship can be accounted for and a more realistic result is obtained.

To obtain fuel weight, it is necessary to assume a typical flight operating time t. Then

$$W_f=CtP \qquad (1)$$

where

W_f = weight of fuel, pounds
C = fuel consumption factor, pounds per kilowatt-hour
P = system losses, kilowatts

As an example, consider the following aircraft electric system: one 30-kw source (equivalent); 100-foot cable (equivalent); 3-kw conversion equipment (rating); 10-hour typical flight period; and C (given fuel and engine) = 0.9 pound per kilowatt-hour.

	30-Volt D-C	120-Volt D-C	120/208-Volt A-C
Losses, kilowatts, P	5.30	5.30	8.55
W_f	47.7	47.7	76.9

For a 4-machine system, these weights would be 191, 191, and 308 pounds respectively, as compared to a total system weight, including load equipment of about 2,000 to 3,000 pounds. Using very approximate estimates of the parameters for evaluation of medium to large aircraft, the fuel weight caused by losses is found to be of the order of 10 or 15 per cent of total electric system weight including load equipment.

Application to a Typical Medium Bomber

The method for computing fuel weight can now be applied to a typical large airplane and, if generation and cable weight are also considered, a system comparison can be made. The system generating capacity is taken as 72 kw and the weights for the three systems are shown in Table II. This indicates that the higher voltage d-c system is lighter in weight than the 3-phase a-c system of the same capacity by about 100 pounds for a medium bomber aircraft. This difference is not significant, but the parts of the system which differ in weight for the various methods of supply are shown in Table II.

Load Equipment Weight Comparison

A comparison of electric systems by weight can be somewhat misleading if the entire systems, including load or utilization devices, are not included in the analysis. Weight advantages or disadvantages which may be hidden if load equipment is not considered are those for control and protection. These usually appear in the generating end of a power system and increase in weight with the degree of control desired. One might attempt to justify, by an apparent weight advantage, a reduction in requirements on frequency and voltage regulation. Considering only the generation and transmission sections of the system, an advantage could be shown. However, all load equipment on such a system would have to be designed and constructed to withstand and be operable on voltages and frequencies which vary over a wide range. The design of load

Table II. Comparisons for a Medium Bomber for 30-Volt D-C, 120-Volt D-C, and 120/208-Volt 3-Phase A-C Systems

System Component	Approximate Weights, Pounds		
	30-Volt D-C System	120-Volt D-C System	120/208-Volt 3-Phase A-C System
Generation equipment, includes constant-speed drives for a-c system	750	800	850
Cable	1,100	950	950
Fuel necessary to supply losses, using the *C* and *t* of equation 1	120	120	190
Miscellaneous weight	150	150	150
Totals	2,120	2,020	2,140

devices would have to be altered for operation at a lower frequency. Certain classes of equipment sensitive to voltage or frequency for proper operation require a regulator or converter unit built into or accompanying the equipment. Because the number of load units is large compared to the source and control units, total system weight (source to and including load) is substantially greater than that of the generating system alone.

The effect on total system weight of the choice of protection equipment is not as readily apparent but still exists. For example, if overcurrent protection is used throughout the system, the generators must have substantial current capability beyond their rating which may give rise to excessive recovery voltages, requiring adequate insulation in load equipment. The choice of protection influences the number of sources, transmission channels, and spare load devices required. The amount of physical protection, duct work, or armored bus boxes, is related to the type of electric protective devices chosen. The use or omission of overvoltage or underfrequency relays for protection may influence the design and weight of load equipment.

The inherent characteristics of certain devices may give them a weight advantage over other types of devices for performing specific functions. For example, a smaller and lighter d-c relay can be constructed than an a-c relay to perform the same function. High-starting-torque, heavy, intermittent-duty motor loads are also more economically served by d-c than by a-c machines. For continuous duty and fixed speed motors, a-c supply offers an advantage.

To measure the weight saving possible by applying the appropriate power supply for each function, it is necessary to have information on the characteristics of devices, and the ratio of weight of the devices to the total system weight. The latter is of importance because the final choice of system supply should be a single supply, with a minimum of conversion and load equipment requiring a type of supply different from the primary supply. A difficulty in obtaining weight figures for load devices is that only the weight affected by electrical considerations is of interest. The weight of equipment, from an airframe manufacturer's weight lists, does not have too much significance since an appreciable part of the weight of each item is that of structural parts, covers, housings, brackets, and the mechanical load devices. Cooling requirements affect the weight of equipment as temperatures increase and as more items of equipment are put into a given volume. Fans, cooling vanes, and in some cases even package refrigerating systems become part of the load device. It is still possible to consider the load functions to be performed and to compute, approximately, the weight of the device necessary to perform the task for a given method of supply.

The first step in such an analysis is to procure load weight data for the airplanes which are to be studied. These data apply to a particular primary system, say a 30-volt d-c system, and certain general considerations indicate approximately what the weights will be for the other two methods of supply. It is the means of finding load equipment weight for a system when the weight is known for another system.

Conversion of Load Weight Data From One System to Another

Loads on aircraft may be classified as:

1. Heating and deicing.
2. Electronic equipment, radio and radar.
3. Motors and actuators.
4. Conversion equipment, defined as a load.
5. Miscellaneous, relays, lighting.

The only items for which weight is significantly affected by the type of supply are items 3 and 4.

Motors which perform many functions constitute a large proportion of the electric load on all kinds of modern aircraft. Some of these are used for long enough periods so that they are rated for continuous duty. Others are rated for intermittent duty when used only for short periods. For either application, the motor must be able to perform without overheating. Another requirement is that adequate torque must be supplied to start the load. A further factor to be considered in motor application is that the heavy starting currents of relatively large motors may cause a depression of the system voltage. The latter can give very undesirable effects such as faulty operation of protective devices or instability of control devices.

D-c motors make available very high starting torques. For series, shunt, and cumulative-compound machines, current is approximately proportional to torque. The continuous power rating of a d-c motor depends on the temperature limitation, but for short periods of time the motor can deliver higher output. The starting period is usually so short that the electric constants and mechanical design determine the maximum power that can be delivered; generally, starting or short-time torques available range upward from five times the continuous torque rating of the machine.

A-c induction motors have a degree of flexibility in that the design may be adapted to the requirement. Motors for continuous duty have a typical starting torque up to about 150 per cent of rated torque. Motors for intermittent duty have typical starting torques of three times rated, achieved with some sacrifice in rated-load efficiency. The selection of an a-c machine may be made on the basis of required starting torque. This may mean that a much larger machine is used than would be required on the basis of full-speed power requirements. Some loads require a high starting torque to overcome static friction, but others such as pumps, fans, and generators require little starting torque. In the latter devices, load increases with speed and adequate accelerating torque is readily obtained.

A comparison of d-c and a-c motor weights for both continuous and intermittent duty may be made on the basis of typical designs. Such a comparison does not account for all special requirements such as an unusually high starting torque for which an a-c machine may be impractical. In most cases, the starting current required by any one motor should not dangerously reduce system voltage, so this factor will be neglected in weight comparisons.

Bergslien, Finison, and Stratton[1] have compiled data for weights of different kinds of motors as a function of torque. These data may be closely approximated for continuous duty motors as follows:

For d-c open-type motors

Table III. Load Equipment Weight Analysis for Typical Aircraft

	Estimated Weights, Pounds		
Type and Equipment	30-Volt D-C System	120-Volt D-C System	120/208-Volt 400-Cycle A-C System
Medium Bomber			
Heating and deicing*	300	300	300
Electronics*	1,600	1,600	1,600
Lighting*	100	100	100
Relays	100	120†	150‡
Motors and actuators			
Continuous duty	750§	750§	450‖
Intermittent duty	500§	500§	660¶
Conversion equipment	425	425	100**
Miscellaneous devices	20	20	20
Totals	3,795	3,815	3,380
Cargo			
Heating and deicing*	20	20	20
Electronics*	850	850	850
Lighting*	100	100	100
Relays	120††	145†	180‡
Motors and actuators			
Continuous duty	475§	475§	280‖
Intermittent duty	400§	400§	520¶
Conversion equipment	90§	90§	60**
Miscellaneous devices	20	20	20
Totals	2,075	2,100	2,030
Fighter-Interceptor			
Heating and deicing*	100	100	100
Electronics*	900	900	900
Lighting*	20	20	20
Relays	30††	40†	50‡
Motors and actuators			
Continuous duty	100§	100§	60‖
Intermittent duty	200§	200§	260¶
Conversion equipment	170§	170§	80**
Miscellaneous devices	20	20	20
Totals	1,540	1,550	1,490
Trainer			
Heating and deicing*	25	25	25
Electronics*	250	250	250
Lighting*	20	20	20
Relays	30††	40†	50‡
Motors and actuators			
Continuous duty	65§	65§	40‖
Intermittent duty	55§	55§	70¶
Conversion equipment	70§	70§	10**
Miscellaneous devices	20	20	20
Totals	535	545	485

*For a given power output, voltage level or choice of a-c or d-c has negligible effect on weight.

†Allowance for longer creepage distances and contact separation.

‡Allowance for either rectifiers or a-c solenoid weight.

§For given power, voltage level has negligible effect on weight.

‖The weight of the equivalent a-c motor is approximately 0.6 times the weight of d-c equipment.

¶The weight of the equivalent a-c motor is approximately 1.3 times the weight of d-c equipment.

**Assume 10 per cent of installed a-c kilowatt capacity must be converted to d-c by transformer-rectifiers at 10 pounds per kilowatt.

††Estimated weight of existing equipment.

$$\text{Wt}=11.5T^{0.56}\text{ pounds} \tag{2}$$

For a-c 400-cycle 3-phase open-type motors

$$\text{Wt}=6.7T^{0.56}\text{ pounds} \tag{3}$$

where

T = rated torque, pound-feet
Wt = weight

In comparing motors for operation at the same speed, the torque is related to the power output by a constant. Therefore, the weight ratio of d-c to a-c motors chosen on the basis of continuous operation for the same power requirement is

$$\frac{\text{Wt of d-c motor}}{\text{Wt of a-c motor}}=\frac{11.5}{6.7}=1.7$$

Consider high-starting-torque a-c motors designed to give three times the torque that could be maintained at rated speed, and d-c motors to give 12 times the torque at rated speed (the latter value is high, but not an upper limit). Then to compare a-c and d-c motors, the quotient of equations 2 and 3 is taken

$$\frac{\text{Wt of d-c motor}}{\text{Wt of a-c motor}}=1.7\left(\frac{T_{\text{d-c}}}{T_{\text{a-c}}}\right)^{0.56} \tag{4}$$

To obtain the same starting torque, substitute ratios for $T_{\text{d-c}}$ and $T_{\text{a-c}}$ to obtain

$$\frac{\text{Wt of d-c motor}}{\text{Wt of a-c motor}}=\frac{11.5}{6.7}\,\frac{(T_{s/12})^{0.56}}{(T_{s/3})^{0.56}}$$
$$=1.7(1/4)^{0.56}=0.78$$

where T_s is a certain starting torque.

The resulting ratio of weights shows an advantage which might be obtained in using d-c motors for high-torque, very-short-time operation. Presuming that either the smallest d-c motor or the smallest a-c motor might be applied to a mechanical load, the choice between the two would depend upon the duty cycle and speed required. The weight ratios of d-c to a-c motors for various applications should fall in the range 0.78 to 1.7, but are occasionally outside. Since the average is considerably greater than 1, the total motor weight of an aircraft should be heavier for a d-c electric system than for an a-c system, unless there is a large proportion of motor loads requiring a high torque for a very short period.

All aircraft require some conversion equipment regardless of their primary system since certain equipment items cannot be standardized at one voltage. For example, all electronic and radar equipment requires both 30-volt d-c and 115-volt a-c power. Conversion equipment weight can be computed by assuming that transformer rectifiers weigh approximately 10 pounds per kilowatt, and that the inverter weight is given by the empirical equation

$$\text{Wt}=44\log_{10}(\text{kva})+38\text{ pounds} \tag{5}$$

which has been deduced from weight data of several manufacturers.

Load Weight Comparisons for Particular Aircraft

Table III gives a load equipment weight comparison for loads designed to operate from different means of supply for four classes of military aircraft: bomber, cargo, fighter-interceptor, and trainer. The results indicate that the differences in weight of total load equipment, when 28 volts d-c or 120/208 volts (3-phase, 400-cycle) a-c are used, are negligible compared to the masking effect of weight not dependent on electrical considerations, probable errors in estimating unknown weights, and the effect of any hydraulic and pneumatic power supply. The latter factor prevents evaluation of the full advantage of various electric systems.

It may be noted that the weights are different for the higher voltage d-c and 3-phase a-c systems. The a-c conversion equipment is lighter than the d-c equipment for the same converted power. For the larger aircraft, this weight difference can be several hundred pounds. The differences in motor weights are also significant. The intermittent-duty motor load weight can favor the d-c system by as much as 150 pounds while the continuous duty motor weight can favor the a-c system by more than 200 pounds. Thus, if for a large airplane the assumption is made that the total motor weight (intermittent plus continuous) is the same for both systems, the a-c system would still be somewhat lighter because of conversion equipment.

Other Bases of Comparison

The reliability of components and systems depends on the complexity and ruggedness of the device or system in question. Both mechanical and electrical factors must be considered. An approximate measure of either is the total number of parts, or the number of parts which function in series (failure of one causing failure of the device or system), and the number which function in parallel (failure of one not necessarily causing failure of the device or system). Modes of operation and environment also affect the reliability of a given item of equipment. Operating speed, temperature and cooling provisions, vibration, shock, acceleration, and continued exposure to severe climate will alter the life expectancy of a device which meets a standard general specification.

If data are obtainable on the performance history of components in the device or system being studied, such that good estimates of the number of initial, wear-out, and chance failures of an item over a given period of time can be made, the probability of failure of the device or system can be approximated. It would be expected that equally complex systems could be brought to the same degree of reliability by development effort on the less reliable elements of the system. This introduces a time factor, development time, into the evaluation of equipment for the immediate future which becomes less significant for long range comparisons.

A recent paper[2] mentions four methods of improving combat aircraft reliability. These are:

1. Selective inspection of components.
2. Reduction in number of components.
3. Conservative design.
4. Parallel duplication.

Each of these methods must be compared with the others in order to make a choice which improves reliability with minimum increase in weight and minimum decrease in performance capability. For an equitable comparison of systems, it must be assumed that all systems to be compared are made up of components of equal quality and of conservative design and that an equal duplication of critical components is provided. Remaining to be evaluated are the number of components and the probability of failure of the components.

A d-c system has fewer components than the controlled frequency a-c system. In the generating or source section of the system, the a-c system requires additional control and metering equipment, namely, real and reactive load division circuits, governor as well as voltage regulator circuits, an exciter and exciter protection and control relays, and possibility of underfrequency protection. A major item is the generator drive unit. Without presenting a discussion of the reliability or degree of development of constant speed drives, mechanical, hydraulic, or pneumatic, the fact remains that an additional item of equipment is required over that needed for a d-c system. The presence of an additional unit, which consists of several mechanical components in "series," must be considered a detriment when compared to a system which does not require such a device on a basis of reliability.

The magnitude of load which must be supplied with alternating current will affect the reliability of the total system. Transformer-rectifiers being static devices can probably be made more reliable than the best inverter. The problem then depends on the degree of importance of the a-c loads to flight or to the mission. The reliability of supply to a-c loads on an airplane with a primary d-c system may be less or equal to that on an airplane with a primary a-c supply. In other respects, all systems, a-c, d-c, at 30 volts or higher, are subject to the same environment, the same functional loading, the same degree of conservatism in design, and care in installation. The variations possible in these factors may outweigh the shortcomings of the present development of components for higher voltage systems.

From ground operation through airborne operation, the a-c system requires more operating attention. On the ground when external power is connected to the airplane, the operator on a d-c equipped airplane need check only voltage. On an a-c equipped plane, in addition to checking voltage, frequency and phase sequence must be checked. If external power of improper phase sequence is connected to the bus, motors will run in the wrong direction, with consequent damage to limit switches and to stops on actuators, and also malfunction of and damage to other rotating equipment.

With engines running, it is necessary only to close the generator switches to get d-c generators on the system. During operation, load meters are compared and voltage rheostats adjusted for proper load sharing among generators. To get alternators on the line, the exciter control relay must first be closed. If one or more alternators are already on the system, the operator must then observe the synchronizing lamps and adjust the frequency rheostat until the machine being put on is operating at very nearly the same frequency as the system. During one of the momentary periods when the alternator and system voltage are in phase, the alternator circuit breaker is closed. During flight the kilowatt and kilovar meters must be monitored and real power and reactive power load sharing maintained by adjusting the frequency and voltage rheostats respectively.

Ease of maintenance is a difficult feature to evaluate quantitatively. In general, maintenance of the alternator drive presents a problem not found on the d-c system. The responsibility for drive maintenance will probably not be the responsibility of electrical specialists, but the line crew and electrical specialists will have to diagnose a drive failure from the observed performance of the electric system. The installation and the adjustment of 3-phase circuits are more difficult for maintenance personnel of limited training, which is another problem. The various phases must be color-coded or marked, and equipment built for line-to-neutral voltages must not be connected line to line. The additional control and protection equipment of the a-c system may result in further maintenance difficulty, given the same degree of reliability of components for a-c and d-c systems.

Conclusions

1. Considering the system and not the loads, the 120-volt d-c system weighs slightly less than either the 30-volt d-c system or the 120/208-volt 3 phase a-c system. The a-c system tends to be heavier because of the weight of the constant-speed drive and the greater amount of fuel needed to supply the system losses, while the 30-volt d-c system tends to be heavier because of the larger cable weight.

2. Load weight comparisons, if disclosing an advantage for any method of supply, give a slight weight advantage to the a-c supply for the aircraft studied. The difference, though, is not very significant and it is dependent on the weight of conversion equipment. If future aircraft were to have idealized systems with no conversion equipment, a survey of the load equipment would not indicate as great an advantage for the a-c supply.

3. Aircraft d-c systems are more reliable than the a-c systems because of the extra components used in the latter, all other components having identical reliability.

References

1. Economic Factors for Aircraft Electric Systems, R. M. Bergslien, H. J. Finison, L. J. Stratton. *AIEE Transactions*, vol. 73, pt. II, Nov. 1954, pp. 270–79.

2. A Designer's View of Combat Aircraft Reliability, I. H. Driggs, J. E. Forry. *Journal*, Society of Automotive Engineers, New York, N. Y., vol. 61, no. 2, Feb. 1953.

◆

Discussion

D. W. Exner (Boeing Airplane Company, Seattle, Wash.): Let us consider the validity of the assumptions made by the authors. The most obvious error appearing in Table I lies in the assumption of one per-unit load for calculation of losses. We do not design primary power systems to operate at loadings equal to the installed capacity. In fact military specifications require 100 per cent initial margin. Although most of the efficiencies used are already too high for present equipment at full load, they must be reduced drastically for part-load operation. For a long-range airplane the conditions for average cruise load will generally be used in all but the most detailed studies.

In basing their tabulation of comparative system losses and weights on equal generating capacities (stated as 72 kw) the authors do not present a true picture. System selections are actually made on the basis of work loads performed, and must therefore reflect the effects of magnitude and efficiency of any power conversions required between generation sources and load sinks. Because of the poor efficiency of rotary inverters as compared with transformer-rectifiers, especially at part load, 8 to 10 per cent, or greater, generator capacity must be added in a typical medium bomber or transport primary d-c system. This increases both the system losses in Table I and the weights in Table II for the d-c system. In the practical case, of course, we must deal with available equipment ratings, but the authors furnish no clues as to their selections, so that adequate appraisal is impossible.

In showing the relative losses of the three systems, the authors make the error of not listing the losses of the d-c system wiring but listing as a direct loss in the a-c system that supposedly caused by the power factor effect. The correct way of handling the a-c wiring loss in this study would have been to use the inverse of the per-unit efficiency in the first bracket term and multiply it by a per-unit value greater than 1.0 to show the increased current resulting from the power factor. This is because high-voltage system wiring is chiefly limited by wire rating instead of voltage drop. Thus the wire weight shown in Table II should be increased for the a-c system above the weight shown for the 115-volt d-c system. It would be interesting to know where the authors obtained a figure of 0.95 for the power factor of a primary a-c system of this size. Actual figures on typical systems are around 0.85. When these factors are taken into account, together with realistic efficiencies and the increase in d-c system capacity required by the higher conversion losses, the 28-volt system loss total is by far the highest, the a-c system loss is next, and the 115-volt d-c system loss slightly less. The difference in losses of the two high-voltage systems will not usually be very significant.

The wiring weight differences shown between the 28-volt system and the two high-voltage systems are highly questionable. Assuming a bomber having 1,100 pounds of wire for a 28-volt system, the high-voltage system wire weights can be expected to be at least 400 pounds lighter. After all, is not this wire weight saving one of the major reasons for going to higher voltage systems?

In their introductory paragraphs the authors state that reliability is secondary to weight. While admitting that the term "reliability" is hard to define numerically, let me assure the authors that the airframe companies attach first importance to reliability, though keeping due emphasis on light weight. Other factors, not mentioned, which influence the designer's choice of a system are the effect son airplane performance, such as drag, effect on engine thrust, approach speed in jet aircraft, and landing distance. Also considered are the heat rejection problems, including the important item of where the losses are located in relation to the means for dissipating them.

Attention is also directed to the misstatement that torque of a series d-c motor is approximately proportional to current. The paper cannot be considered as an effective representation of system selection criteria or as an accurate listing of typical weights.

Lee R. Larson (United States Naval Research Laboratory, Washington, D.C.): This timely paper is of general interest to the industry. However, since very few of the basic assumptions are listed, serious question is raised regarding the validity of the conclusions. A request is made that the authors define fully the basic conditions applying to the material presented. With this information, the reader would be in a position to apply the techniques presented to his particular problem. In general, the specific selection of an electric system must be made by studying the particular aircraft in question.

H. L. Garbarino, A. K. Hawkes, and **J. A. Granath:** To a great degree, the subject of the paper is controversial. In the past 15 years there have been spirited arguments on the selection of aircraft electric systems, and arguments regarding prime considerations in philosophy of design. To meet pressing needs, decisions on policy have been made in the aircraft industry and by government agencies. Important decisions are often made on the basis of incomplete data. While this is a practical necessity from the standpoint of national welfare, it is regrettable that there is often great feeling against a review of past decisions in the light of new information. Engineering habits and millions of contract dollars are powerful arguments for conservatism.

In view of these factors, the authors believe that one of their most important qualifications for undertaking such a study is the fact that they are associated neither with the airframe manufacturers nor the equipment manufacturers. However, an attendant disadvantage is that detailed data are sometimes difficult to obtain, even with the aid of the manufacturers, which is usually given generously. Therefore, to a degree it has been necessary to use estimates and approximations.

In comparisons of the type undertaken in the paper, it is necessary to make some compromise between actual electric systems which could be assembled in a short time, and those ideal systems which ignore the problem of equipment availability. It is

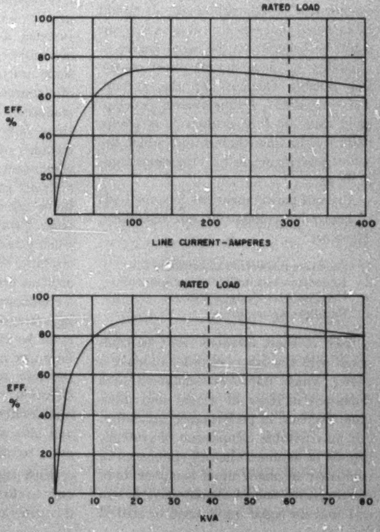


Fig. 1. Typical efficiency curves. Top—Westinghouse type D-30 aircraft d-c generator. Bottom—Westinghouse alternator 8QL-40A

not justifiable to study long-term problems limited to present production or present designs. It might be noted, however, that there is a rather regrettable tendency for the industry to be guided in long-term design by currently available equipment. A paper of this kind would be valueless under such a restriction.

With regard to the cable weight values in Table II, the authors agree that the principal reason for substituting a 120-volt d-c electric system for a 30-volt system is the cable weight saving. However, when a considerable portion of the cable weight is of minimum wire size at 30 volts, the weight saving effected by increasing voltage is lessened. For the aircraft considered in Table II, approximately 40 per cent of the cable weight is at minimum wire size at 30 volts. It is true that the cable weight difference between the two d-c systems can be as high as 400 pounds, but after studying several aircraft employing 30-volt d-c systems with about 1,000 pounds of cable, the authors believe the difference will usually be less.

Although primary power systems are not designed to operate at loads equal to installed capacity, there has been, on post-World War II military aircraft, a tendency for electric load equipment to be added after the design of the primary system has been fixed. Because of the load growth which occurs during preproduction development, and the operating life of the aircraft, the 100-per-cent initial margin of capacity is heavily taxed. From load analysis charts for several long-range aircraft (operating in 1952–53) it would appear that a rough figure for average cruise load is 70 per cent of installed primary capacity. Curves of efficiency versus load for machines and devices in general usually show a high efficiency over a wide range of loads, the maximum occurring at something less than rated load. Efficiency curves for an a-c and a d-c generator obtained from manufacturers' data are offered as examples. (See Fig. 1.) On this basis the authors cannot agree with Mr. Exner that efficiencies must be drastically reduced for typical in-flight load.

The authors believe that the problems of reliability and weight of military aircraft are much more closely related than Mr. Exner indicates, particularly when it is realized that other factors such as over-all simplicity, number of components, and essentially weak parts are considered. The problem is much too important to dismiss with a statement of intent. This view is somewhat emphasized by the general excitement during the Korean War on comparisons of various aircraft, their performance, and their weights.

It would be interesting to see how a general analysis of this type might incorporate quantitatively the effect on the electric system of drag, thrust, approach speed, landing distance, etc. We hope that Mr. Exner's next treatise will cover these matters. With regard to the statement concerning series d-c motors, the authors are aware that torque is ideally proportional to the square of current. However, in practical motors, the effect of saturation of the magnetic materials at operating densities, and the effect of armature reaction, are such that a better approximation is obtained with an exponent of unity rather than two, although some intermediate value might be best.

Experimental Determination of 400-Cycle Impedance of Wire in Aircraft Power Distribution Circuits

J. D. ANDREW
ASSOCIATE MEMBER AIEE

The Problem and its Background

THE measurement of impedance of power distribution circuits is a continuing problem as long as materials, installation methods, and ultimate use of such circuits are constantly undergoing changes. In contrast to transmission lines, power distribution circuits are relatively short and operate at utilization voltages. This simplifies the impedance measuring problem because the effects of capacitance and leakage are negligible at power frequencies. Consequently, the literature shows that the analysis of the problem has been thorough and measurement techniques have been painstaking; however, the measurements have usually been made at 60 cycles and, in the case of 3-phase measurements a neutral wire or earth return has been used. The 3-phase 4-wire electric systems rapidly being developed for aircraft in this country are operated at 400 cycles, and the neutral return is made through the aircraft structure, particularly the conducting aluminum skin. For increased physical reliability, the wire used in aircraft distribution circuits has a different construction than that commonly used in other applications. These differences are sufficient to hinder the use, even by extrapolation, of the results of much research on the general problem of wire impedance measurement.

The need for accurate and extensive information concerning all components of aircraft electric systems becomes more acute as designers begin to recognize that these systems are highly integrated. Their performance depends not only upon the individual electrical properties of each component but upon composite system characteristics. The impedance of the distribution wiring is a factor of major importance when a system is subjected to faults or other heavy transient disturbances. Its effect on steady-state operating conditions is becoming more important as aircraft grow in size.

An excellent approach to the computation of aircraft wire impedance has been made in references 1 and 2, which have provided great assistance to many designers. It is readily evident from these writings that an analytical effort to determine wire impedance is hampered by the mathematical difficulties which arise in calculating the distribution of return current through the aircraft skin. Experimental measurements of aircraft wire impedance have been limited to the measurement of average zero-sequence impedance, and these results have been used to supplement analytical calculations or to substantiate necessary simplifying assumptions. The approach to the problem which was used in obtaining the results reported here has been totally experimental; hence, rigorous analytical explanations of the results are lacking.

The Impedance Function in 3-Phase Circuits in Aircraft

USE OF SEQUENCE IMPEDANCES

The impedance of a single-phase 2-wire distribution circuit is classically defined as the ratio of voltage drop per unit length of the line to the line current, and this is

---

Paper 54-374, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing August 13, 1954.

J. D. ANDREW is with the Douglas Aircraft Corporation, Santa Monica, Calif.

Sincere appreciation is expressed to D. P. Atchison and B. H. Owens of the Douglas Aircraft Corporation for their invaluable assistance in conducting these tests and evaluating the results.

Table I. Characteristics of Wire Used in Tests

Size of Cable	Diameter of Conductor, Inches	Diameter of Finished Cable, Inches	D-C Resistance, Ohms per 1,000 Feet
AN-0	0.401	0.505	0.098
AN-2	0.330	0.425	0.157
AN-4	0.253	0.355	0.255
AN-6	0.217	0.312	0.405
AN-10	0.128	0.203	1.013
AN-14	0.076	0.144	2.516
AN-18	0.051	0.110	5.568
AL-0	0.410	0.535	0.154
AL-2	0.321	0.423	0.249

Notes:

1. Applicable government specifications for aircraft wire are MIL-W-5086[7] for copper wire, designated as *AN*, and MIL-W-7072[8] for aluminum wire, designated as *AL*.

2. These samples are representative of the production stock of one aircraft manufacturer and conform to more rigid specifications than the applicable government specifications.

generally a complex number. The real part of this number is the line resistance and the imaginary part is its reactance. The extension of this concept to 3-phase lines is easily accomplished provided that the assumption of complete symmetry is justified. A further extension can be made by resorting to the methods of symmetrical components and the use of positive-, negative-, and zero-sequence impedance values. For a completely symmetrical line, the positive-, and negative-sequence impedances are equal and, together with the zero-sequence impedance, they will define the behavior of the line under all operating conditions. It must be observed that this rather simple extension of a single-phase concept to a 3-phase 4-wire system is predicated on line-to-line and line-to-neutral symmetry, and that in the unsymmetrical case the use of only one value of positive-sequence impedance and one value of zero-sequence impedance is without meaning unless qualified as average values. Chapter 8 of reference 3 shows clearly that nine distinct impedance values, three for each phase, are required to completely predict the behavior of a 3-phase 4-wire line with no symmetry. Experimental determination of these values can be made by measuring the impedance of each conductor, first with only positive-sequence currents flowing, then with only negative-sequence currents flowing, and finally with only zero-sequence currents flowing. The method of accomplishing this is described in Appendix I.

NATURE OF REAL AND IMAGINARY TERMS OF IMPEDANCE FUNCTION

It is a common experience in the measurement of wire impedance that three identical conductors will exhibit widely differing values of positive-sequence resistance when they are used as a 3-phase distribution circuit. The reason for these variations is the dissymmetry of the line, and an analysis of this effect is made in reference 4. It is there shown that the inductive reactance of a conductor of a 3-phase line is a complex number in itself and that resistance measurements of such a conductor will include the real part of the inductive reactance as well as the "ohmic" or single-phase resistance of the conductor. Some of the common configurations of aircraft wiring are sufficiently unsymmetrical to make this induced component of conductor resistance nearly equal to single-phase resistance. If the neutral return of a 3-phase circuit consists of a conductor, then zero-sequence impedance measurements of conductor resistance will include only the single-phase resistance of that conductor even though the line is unsymmetrical; however, the use of the aircraft skin as a return path for zero-sequence currents introduces a very significant change in the measured resistance of this same conductor. It should be emphasized, therefore, that a resistance value obtained from wire tables is not at all indicative of the real term of the impedance of that same wire in a 3-phase circuit.

Table II. Positive-Sequence Impedance for Equilateral Configuration

Wire Size	Impedance, Ohms per 1,000 Fee Minimum Elevation	 6-Inch Elevation
AN-0	0.106+*j*0.182	0.101+*j*0.185
AN-2	0.165+*j*0.192	0.158+*j*0.195
AN-4	0.262+*j*0.201	0.256+*j*0.201
AN-6	0.413+*j*0.218	0.409+*j*0.220
AN-10	1.018+*j*0.227	1.014+*j*0.227
AN-14	2.507+*j*0.260	2.503+*j*0.263
AN-18	5.531+*j*0.286	5.529+*j*0.286
AL-0	0.165+*j*0.192	0.158+*j*0.196
AL-2	0.255+*j*0.198	0.253+*j*0.206

EFFECT OF USING A FLAT SHEET OR SKIN AS A NEUTRAL RETURN

Lack of accurate knowledge regarding the distribution of zero-sequence currents in the flat sheet or skin return of aircraft circuits renders mathematical analysis extremely difficult if the lumped constants of circuit theory are relied upon. An approach using field theory will probably express the situation more precisely; however, insight can be gained into the behavior of this type of circuit element by considering the skin to be a finite number of filamentary type conductors lying in a flat plane. This approach is developed in reference 2, from which it may be deduced

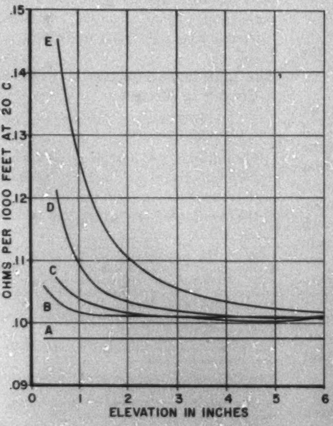
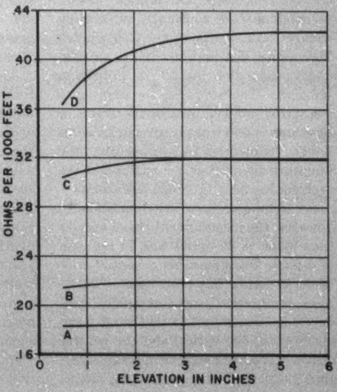


Fig. 1 (left). Positive-sequence resistance of AN-0 wire

A—D-c resistance
B—Tight equilateral configuration
C—Flat group, minimum spacing
D—Flat group, 1-inch spacing
E—Flat group, 2-inch spacing

Fig. 2 (right). Positive-sequence reactance of AN-0 wire

A—Tight equilateral configuration
B—Flat group, minimum spacing
C—Flat group, 1-inch spacing
D—Flat group, 2-inch spacing



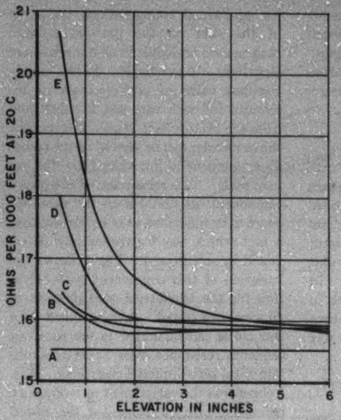


Fig. 3 (left). Positive-sequence resistance of AN-2 wire

Curves A–E same as for Fig. 1

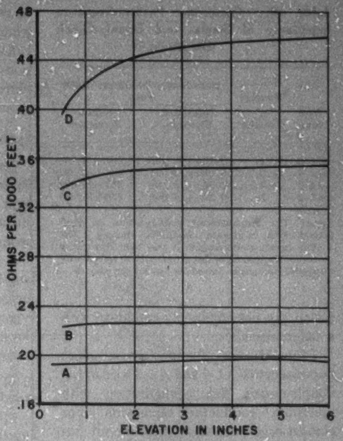


Fig. 4 (right). Positive-sequence reactance of AN-2 wire

Curves A–D same as for Fig. 2

that because the filamentary skin currents are in various time-phase relationships to the conductor currents, the inductive reactance measured along the conductors will be a complex number. As described in the foregoing, this will cause the measured or apparent resistance of the conductors to be widely different from the single-phase resistance of these conductors. Most of the commonly used configurations of aircraft distribution wiring have a high degree of line-to-neutral symmetry and, as a consequence, little variation would be expected between zero-sequence impedances of the three line conductors.

In addition to its effect as a zero-sequence current return path, the aircraft skin will also alter the positive- and negative-sequence impedances of distribution circuits. Inherent line-to-neutral symmetry of most configurations greatly reduces the magnitude of this change; however, the presence of the skin near a distribution circuit will distort the magnetic field surrounding the conductors and absorb power from them because of the induced currents which will circulate in the skin. These effects, although external to the line conductors, will be reflected as an increase in average circuit resistance and a decrease in average reactance.

Impedance Value for Selected Measurement Conditions

The most important independent variables in this type of measurement are:

1. Wire size and conductor material.
2. Conductor configuration.
3. Elevation above skin section.
4. Current balance.

The combinations of these variables are numerically limitless, and it became necessary to select specific sets of conditions for investigation. Table I shows the nine wire sizes used, along with the d-c resistances of these samples. All sizes were tested in the tight equilateral configuration. Sizes *AN-0* and *AN-2* were tested in the flat configurations and the center-to-center spacing was varied from minimum to 2 inches. Sizes *AL-0* and *AL-2* were tested in a double-circuit tight flat configuration. All resistances were corrected to 20 degrees centigrade to provide a common basis for comparison.

Elevation was measured from the skin surface to the center of the lowest conductor and was varied in six steps from minimum to 6 inches. For reasons described previously, three distinct conditions of current balance were applied, these being positive sequence only, negative sequence only, and zero sequence only. Further details of the measurement method are discussed in Appendix I.

Positive-Sequence Impedance

For a given wire size, configuration, and elevation, a total of six readings was obtained for the positive- and negative-sequence conditions. It would be expected that the average of the positive-sequence impedances would equal the average of the negative-sequence impedances, and the actual measurements showed this to be true within the measurement accuracy. Slight differences between positive- and negative-sequence impedances were averaged out, and this average of the six readings is used as the positive-sequence impedance. Negative-sequence impedance is therefore also equal to this average value. Although individual phase impedances may depart rather seriously from this average figure, it is nevertheless a good comparative criterion of over-all line behavior.

The equilateral configuration, with conductors tightly laced together, is inherently symmetrical and is commonly used in aircraft wiring. The positive-sequence impedances of this configuration for the nine sizes tested are listed in Table II. Two values are given for each size, one obtained at minimum elevation and the other at 6 inches' elevation. Comparison of these values shows that the effect of the skin is not measurable for wire sizes smaller than *AN-6* and that it is very slight for larger sizes.

Flat configurations are also widely used in aircraft, but, as the conductor separation is increased, the effect of the skin becomes more significant. Fig. 1 shows the variation with elevation of the positive-sequence resistance of several of the flat configurations, as well as tight equilateral grouping, for size *AN-0* wire. A flat circuit with 2-inch conductor spacing will exhibit nearly a 50-per-cent increase over its isolated average resistance if routed close to the skin. On the other hand, minimum conductor spacing reduces the increase to less than 10 per cent. Fig. 2 is a plot of reactance variation with elevation of these same circuits. Reactance changes are less marked but still significant. Figs. 3 and 4 are plotted from data for size *AN-2* wire and show a comparable increase in resistance and decrease in reactance as elevation is reduced. Flat configurations with wide conductor sepa-

Table III. Comparing Positive-Sequence Impedance of Single- and Double-Circuit Groups

Wire Size	Number of Circuits	Impedance, Ohms per 1,000 Feet: Minimum Elevation	Impedance, Ohms per 1,000 Feet: 6-Inch Elevation
AL-0	1	$0.165+j0.192$	$0.158+j0.196$
AL-0	2	$0.080+j0.091$	$0.080+j0.092$
AL-2	1	$0.255+j0.198$	$0.253+j0.202$
AL-2	2	$0.128+j0.093$	$0.128+j0.093$

Note: The single-circuit configuration is equilateral with minimum conductor spacing. The double-circuit configuration is one tight flat group on top of another tight flat group with diametrically opposite conductors paired for two of the phases.

ration are not commonly employed with smaller wire sizes.

Table III lists the positive-sequence impedances of *AL-0* and *AL-2* wire in the double-circuit flat configuration with minimum conductor separation. For comparative purposes, single circuit impedances of these same conductors are listed for the equilateral configuration with minimum conductor spacing. The double-circuit configuration was made up by lacing one flat group tightly on top of another and pairing diametrically opposite conductors for two of the phases. The two center conductors were paired for the third phase. This configuration might be viewed as two triangular groups, one being inverted. It is evident that elevation has practically no effect on the positive-sequence impedance of this type of grouping and that reactance is somewhat less than half of the single-circuit values. Many double-circuit configurations have positive-sequence reactances somewhat greater than half of the corresponding single-circuit arrangement.

ZERO-SEQUENCE IMPEDANCE

The zero-sequence impedance of a 3-phase circuit involves a common return path. For a symmetrical line, zero-sequence impedance is defined as the impedance per unit length of one conductor plus three times the impedance per unit length of the common return path. In measuring zero-sequence impedances of typical distribution circuits with a return through a simulated skin section it was learned that:

1. The conductor impedances do not differ greatly from each other.
2. The impedance introduced by the skin appears almost entirely as an increase in conductor impedance. For nearly all test conditions, less than 10 per cent of the zero-sequence voltage drop appeared along the return path.

With only zero-sequence currents flowing, each combination of wire size, configuration, and elevation provided four measured impedances, one for each conductor and one for the skin. To obtain an average zero-sequence impedance, the skin impedance was multiplied by three and added to the average of the three conductor impedances. As stated before, for equilateral configurations the effect of the skin on the positive-sequence resistance is negligible at all but minimum elevations. Accordingly, if the nearly constant value of positive-sequence resistance for each wire size is subtracted from the average zero-sequence resistance the remainder can be viewed as the resistance component introduced by the return path. This remainder, or resistance component attributable to the skin, appears to be independent of conductor size, a fact which was experimentally determined in reference 1. Fig. 5 shows the variation of this component with elevation for the equilateral configuration of all wire sizes tested. A value taken from this curve can be added to the positive-sequence resistance from Table II to obtain total zero-sequence resistance. Fig. 5 also shows the variation of zero-sequence reactance with elevation for the tested sizes. This quantity is not independent of wire size and has no direct relation to the positive-sequence reactances of Table II.

The resistance curve of Fig. 5 should not be interpreted as the measured resistance of the skin return path. To illustrate the measured breakdown of the zero sequence impedance components, Figs. 6 and 7 are curves of conductor resistance and reactance for size *AN-0* wire and for the skin return path, taken with only zero-sequence currents flowing. It will be noted that the return path impedance is

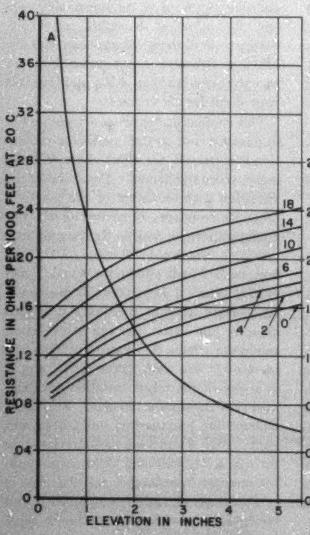


Fig. 5 (left). Zero-sequence impedance for equilateral configuration

A—Increase of zero-sequence resistance over positive-sequence resistance for all wire sizes. Numbered curves show zero-sequence reactance for various wire sizes

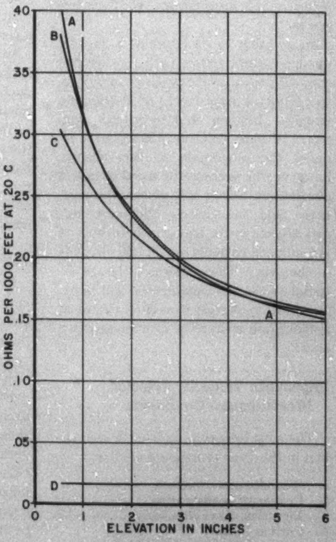
Fig. 6 (right). Zero-sequence resistance of size AN-0 wire (total circuit resistance of one conductor)

A—Equilateral and flat groups with minimum spacing

B—Flat group, 1-inch spacing

C—Flat group, 2-inch spacing

D—Three times measured skin resistance (for all groups)



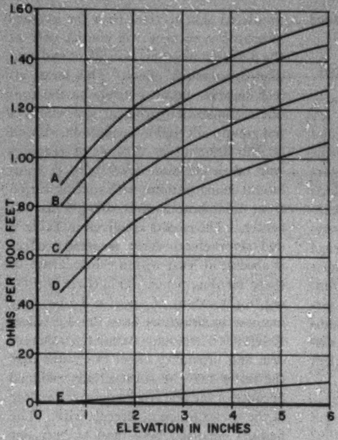


Fig. 7 (left). Conductor and skin components of zero-sequence reactance for size AN-0 wire

A—Tight equilateral configuration
B—Flat group, minimum spacing
C—Flat group, 1-inch spacing
D—Flat group, 2-inch spacing
E—Three times measured skin reactance (for all groups)

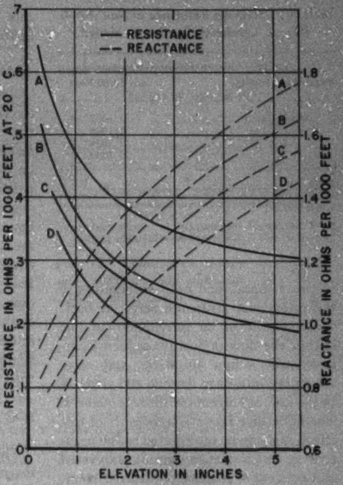


Fig. 8 (right). Zero-sequence impedance of aluminum wire

A—AL-2 wire in tight equilateral group
B—AL-0 wire in tight equilateral group
C—AL-2 wire in double-circuit flat group, minimum spacing
D—AL-0 wire in double-circuit flat group, minimum spacing

relatively small and practically unaffected by elevation. This would indicate that heavy zero-sequence current flow will not cause any significant displacement of the load neutral from the generator neutral as would occur with a neutral wire return. Although Figs. 6 and 7 are plotted for only one wire size, comparable curves for other sizes would show the same values of impedance for the skin. The zero-sequence impedances of some of the flat configurations of size *AN-0* wire also plotted in Figs. 6 and 7 illustrate that increasing conductor separation actually reduces zero-sequence impedance, particularly the reactance component.

The use of double circuits to reduce positive-sequence impedance can be highly efficient, as shown in the preceding but does not achieve a comparable reduction of zero-sequence impedance. Fig. 8 shows the total zero-sequence resistance and reactance of *AL-0* and *AL-2* wires in tight flat 2-circuit configurations and also in tight equilateral configurations. Double-circuit wiring results in only a 25 to 30-per-cent decrease in zero-sequence resistance from the single-circuit resistance and only a 15-per-cent decrease in reactance from the single-circuit value. This further illustrates the relatively large effect of skin return on zero-sequence impedance of power distribution circuits.

Nomenclature

$Z_{a1}, Z_{b1} \ldots$ = impedance of the phase conductor designated by the alphabetical subscript measured under conditions of current flow of the sequence designated by the numerical subscript

I_1, I_2, I_0 = positive-, negative-, and zero-sequence currents

I_a, I_b, I_c = line or test conductor current

V_a, V_b, V_c = fall of potential along test conductors

$R_{a1}, R_{b2}, \ldots$ = real component of $Z_{a1}, Z_{b2}, \ldots$

$X_{a1}, X_{b2}, \ldots$ = imaginary component of $Z_{a1}, Z_{b2}, \ldots$

X_a, X_b, X_c = self-reactance of test conductors

X_{ab}, X_{bc}, X_{ca} = mutual reactance between test conductors

R_a, R_b, R_c = single-phase resistance of test conductors

Dissymmetry in Aircraft Power Distribution Circuits

UNBALANCE FACTORS

The results presented so far are average impedance values and are not necessarily representative of the impedance of any given phase. All nine sequence impedances are required for accurate calculations of line behavior; however, this laborious approach is seldom justified if a reasonable degree of symmetry exists. A figure of merit which, at least to some extent, permits a numerical measure of line symmetry is described on page 247 of reference 3 as the ratio of sequence mutual impedance to sequence self-impedance. Several such ratios are possible, but the most significant are the negative-sequence unbalance factor M_2 and the zero-sequence unbalance factor M_0. The following equations define these factors

$$M_2 = (Z_{a1}+aZ_{b1}+a^2Z_{c1})/(Z_{a2}+Z_{b2}+Z_{c2}) \quad (1)$$

$$M_0 = (Z_{a1}+a^2Z_{b1}+aZ_{c1})/(Z_{a0}+Z_{b0}+Z_{c0}) \quad (2)$$

The quantity a is the usual complex operator for 120-degree displacement. The impedance terms Z_a, Z_b, and Z_c are the individual phase impedances, and the numerical subscripts indicate that these impedances are measured with current flow restricted to only positive or only negative or only zero sequence. For a line with complete symmetry, the positive-sequence impedance of all three

Table IV. Unbalance Factors for AN-0 Wire

Configuration	Conductor Spacing, Inches	Negative-Sequence Unbalance Factors: Minimum Elevation	Negative-Sequence Unbalance Factors: 6-Inch Elevation	Zero-Sequence Unbalance Factors: Minimum Elevation	Zero-Sequence Unbalance Factors: 6-Inch Elevation
Equilateral	minimum	0.002	0.023	0.029	0.003
Flat	minimum	0.272	0.289	0.039	0.024
Flat	3/4	0.207	0.228	0.042	0.027
Flat	1	0.164	0.198	0.049	0.028
Flat	1 1/2	0.112	0.174	0.052	0.030
Flat	2	0.066	0.155	0.052	0.030
Double Flat*	minimum	0.028	0.067	0.033	0.021

*Size AL-0 wire.

Table V. Average Resistance of Size AN-0 Wire as Conductor Spacing is Varied

Conductor Separation, Inches	Resistance, Ohms per 1,000 Feet
Minimum	0.1037
1	0.1038
2	0.1039
5	0.1039

Notes:
1. D-c resistance of this wire is 0.1006 ohm per 1,000 feet.
2. Conductors maintained in equilateral configuration and isolated by removing skin section.

phases is equal, and both numerators in the ratios just given are zero. A basic characteristic of an unsymmetrical line is that the flow of positive-sequence currents will induce negative- and zero-sequence voltages in the line and give rise to the flow of currents of those sequences. The unbalance factor is a measure of this effect, and can be interpreted as the negative or zero-sequence current, expressed as a per cent of positive-sequence current, which would flow if the line were terminated in a 3-phase-to-neutral short circuit of zero impedance. It can be stated in general terms, therefore, that the farther a line departs from symmetry the greater will be the unbalance factors, and a tabulation of these factors for various configurations will present a comparison of the degree of symmetry of these groupings.

Table IV gives negative- and zero-sequence unbalance factors for several configurations at minimum elevation and at 6 inches' elevation for *AN-0* and *AL-0* wire. From this table it is evident that both the tight equilateral and the double-circuit tight flat groupings are highly symmetrical, whereas the single-circuit tight flat group is the most unsymmetrical of those tested. Zero-sequence unbalance factors are generally lower than the corresponding negative-sequence unbalance factors because the average zero-sequence impedance of any configuration is usually much higher than the average negative-sequence impedance. Configurations which have large unbalance factors will produce unbalanced terminal voltages even for balanced loads; however, this effect is not particularly serious except for the longest circuits. Table IV also shows that the presence of the skin affects the unbalance factors and, hence, the symmetry of the line. For all flat configurations, routing the circuit close to the skin will reduce the negative-sequence unbalance factor and, conversely, cause a slight increase in zero-sequence unbalance factor. The effect on tight equilateral configurations is just the opposite; however, this configuration has small unbalance factors at any elevation.

TRANSPOSITION AND RANDOM BUNDLING

Because of their relatively short lengths aircraft power distribution circuits are not ordinarily transposed in any regular fashion. Impedance measurements made on 3-phase circuits with the conductors twisted around each other show a high degree of symmetry compared to flat configurations, but also show little improvement over the already highly symmetrical equilateral configuration. Wires smaller than *AN-6* are ordinarily bundled in random fashion, quite often with other circuits. This usually introduces some unintentional transposition and line dissymmetry does not appear to be a problem with these smaller size circuits.

A-C Resistance of Aircraft Wire

The measurement accuracy applying to the methods used in this investigation is discussed in Appendix II. A higher order of accuracy would be required for a precise evaluation of the factors contributing to the a-c resistance of the wire used in aircraft. Nevertheless, some experimentation was devoted to checking currently available data of this nature.[5]

The measured resistance of a conductor in a 3-phase distribution circuit is seriously affected by factors such as line dissymmetry, current balance, and adjacent metal masses such as the aircraft structure. The influence of line dissymetry and current balance can be eliminated by averaging the resistances measured first with only positive-sequence currents flowing and then with only negative-sequence currents flowing.[4] If the conductors are isolated by removing the simulated skin section, then the average measured resistance will consist only of the d-c resistance plus skin effect and conductor proximity effect. This latter effect depends almost entirely on the conductor separation, whereas skin effect is not greatly affected by separation. Based on this assumption, a series of readings was taken on size *AN-0* wire in equilateral configurations, with spacing varied from minimum to 5 inches, center to center. The results are given in Table V and show clearly that proximity effect is absent, at least within the accuracy of these measurements. Skin effect is present to the extent of about a 3.5-per-cent increase in resistance over the d-c value. Reference 5 lists some calculated values of skin and proximity effect but notes that the latter may be substantially reduced because of the fine stranding and spiraling in the construction of aircraft wire. The findings reported here tend to bear out this contention.

Conclusions

Laboratory measurements of aircraft wire impedance, to the extent conducted by the author and his associates, have confirmed or further substantiated much previous work by others. In addition, it is believed that some of the more obscure aspects of this general problem have been brought into clearer perspective through the ability to assign comparative values to those effects which are difficult to compute. Specific conclusions which were reached are:

1. Unsymmetrical power wiring of 3-phase circuits greatly alters the apparent resistance of the individual phase conductors. Resistance values selected from ordinary wire tables are not directly applicable to the individual phases of unsymmetrical circuits.

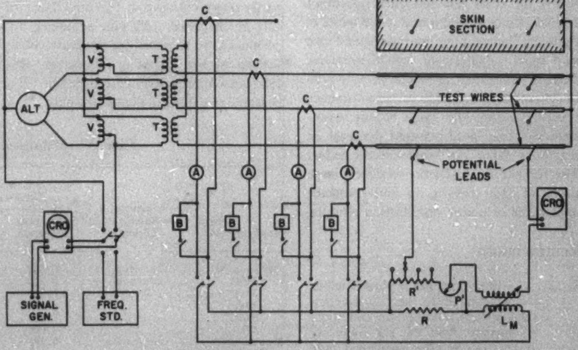


Fig. 9. Circuit diagram for 3-phase impedance measurements

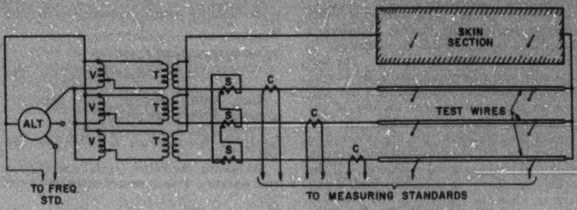


Fig. 10. Alteration of circuit connections for zero-sequence measurements

2. The presence of the aircraft skin or structural members in close proximity to the circuit wiring will affect the positive-sequence impedance of these circuits. Symmetrical circuit arrangements reduce this effect, but the result is always an increase in average resistance and a decrease in average reactance.

3. The use of the aircraft structure or skin as a return path for zero-sequence currents results in an increase in the apparent resistance of the circuit conductors. This increase is comparatively large when the circuit is routed close to the skin but it is practically independent of the conductor size.

4. Single-circuit flat configurations have high impedance if conductor spacing is large and high unbalance factors for any spacing. Such configurations are best avoided for long runs of large wire sizes.

5. Tight equilateral configurations or close random bundling are highly symmetrical and represent an efficient use of conductor material from an impedance standpoint. Double-circuit flat configurations are even more efficient, but the designer must be able to exercise control over the pairing of the phase wires to realize the advantages of this type of grouping.

6. The 400-cycle resistance of large size aircraft wire is not markedly greater than the d-c resistance of this same wire. It appears that this can be attributed to a significant lack of conductor proximity effect.

Appendix I. Impedance Measuring Method

The circuits which were set up to determine the impedance of various distribution lines are shown in Figs. 9 and 10. This is very similar to the method described in reference 6 and is often referred to as the "fall of potential" method. The noninductive resistance standard R and the Brooks mutual inductometer L_M produce voltage drops which are in time phase and quadrature respectively with the current in the line being measured. An adjustable percentage of each of these voltages is then compared with the voltage drop along the line under measurement. When the adjusted voltage output from these standards is equal to the unknown line drop, a null is indicated on an oscilloscope which is connected as a differential voltmeter. The standards are calibrated directly in ohms and microhenrys and, when the readings are multiplied by the ratio of the current transformer C associated with the measured line, they equal the apparent resistance and inductance of that line. The voltage dividers R' and P' connected across the standard R have 1,000 times the resistance of R, and only a negligible part of the measuring circuit current is shunted through them. The extremely high input impedance of the oscilloscope, together with the fact that the differential voltage across it is balanced to a null, prevents the flow of current in the secondary winding of the inductometer L_M. The voltage across this secondary winding, therefore, is truly in time quadrature with the line current. Similarly, no current can circulate in the potential measuring leads connected to the wire under test, hence the impedance of these leads is not a factor in the measurement. The use of an oscilloscope as a differential voltmeter is advantageous because of its high impedance and also because it permits an observer to distinguish between the fundamental 400-cycle supply voltage and any harmonics which may be generated.

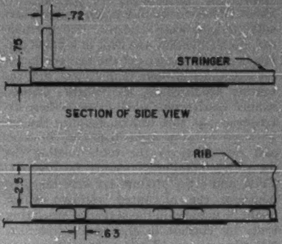


Fig. 12. Construction details of simulated skin section. Skin thickness varies from 0.025 inch at center to 0.078 inch at edges

The standard test length of wire used in these measurements was 8 feet between potential measuring points. Short potential measuring leads were attached to the wires under test by opening the insulation, tightly twisting two wraps of no. 24 brass wire around the conductor, and soldering a quick disconnect device to the brass wire. A single set of potential lead extensions was permanently installed so that observers could rapidly reconnect these extensions to any of the short potential leads of the wires under test. Great caution was exercised in routing the potential lead extensions. Each extension followed a vertical path, perpendicular to the test wires, for a distance of about 4 feet. Each was turned 90 degrees over a nylon pulley and routed

Fig. 11. Measuring standards, frequency monitoring equipment, and null detection oscilloscope

Fig. 13. Skin section inverted for measurements. Supports used to maintain cable elevation

about 12 feet to another nylon pulley where small weights were attached. In this manner, each extension was kept in a plane perpendicular to the axis of the test conductor for a distance of more than 12 feet, and the tension produced by the weights assured against inadvertent small loops in these extensions. At this 12-foot distance the extensions were brought together, twisted, and run down to the table holding the measuring standards, shown in Fig. 11. These and other precautions were taken to minimize the unwanted voltages which might be induced in the potential measuring leads by current in the test conductors. No other source of 400-cycle voltage was operated in the vicinity, so stray voltages did not provide a problem.

To simulate the aircraft structure, a 10 by 4-foot section of *24S* Alclad aluminum sheet was reinforced by stringers and ribs, and then extended to an over-all size of 14 by 6 feet by attaching additional sections of sheet. Details of construction are illustrated in Fig. 12. D-c resistance of the Alclad sheet without reinforcements was measured as 1.082 ohms per 1,000 feet per inch of width. D-c resistance of the complete section was measured as approximately 0.0044 ohm per 1,000 feet. After much experimentation it was determined that the skin section could be inverted as shown in Fig. 13 and the test cables placed on the side opposite the reinforcements without affecting the measurements. This was done to facilitate measurements at low values of elevation. Short potential measuring leads were attached to the skin, 8 feet apart along the centerline, by means of sheet metal screws.

As indicated in Fig. 9, conditions of positive- and negative-sequence current flow were achieved by opening the neutral return and adjusting the variable ratio transformers *V* for equal currents in each line as indicated by ammeters *A*. Four dummy burdens *B* were constructed to match the impedance of the measuring standards. As the measuring standards were switched from one phase to another, a dummy burden was switched in to replace them so that the current transformers *C* presented a constant impedance to each phase.

The alternator *ALT* is a 15-kva 3-phase unit with provisions to adjust frequency and voltage. Its line-to-line voltage was adjusted between 30 and 100 volts during the course of the 3-phase tests. It was adjusted as high as 150 volts line to line for single-phase requirements. Transformers *V* and *T* provided a voltage reduction to less than 10 volts. Current in the test wires was varied from 40 amperes for large wire sizes to 8 amperes for *AN-18* wire. Line current magnitude had no measurable effect on impedance, and values were selected which would stabilize wire temperatures a few degrees above ambient.

Fig. 10 shows the use of three additional current transformers *S* with secondary windings in series to obtain true zero-sequence current flow. If any degree of dissymmetry exists, the classic method of paralleling the test conductors and applying a single-phase voltage will not cause true zero-sequence currents to flow. Unbalanced voltages must be applied to the system, and the current transformers *S* accomplish this by automatically adjusting the line impedances to make the line currents equal and in phase. This method is highly effective because the current transformer impedance is comparable in magnitude to the impedance of the test wires. The fact that only zero-sequence currents are flowing is easily checked after obtaining a null balance in the usual manner for one of the conductors. Without disturbing the settings of the standards, they are switched to another line, but the potential leads remain connected to the original conductor. If the oscilloscope shows no departure from a null balance, then the currents in the two lines are equal and in time phase. The third line can be checked in the same manner.

Frequency was maintained at 400 cycles by continuous monitoring with a signal generator and an oscilloscope. The setting of the signal generator was checked against a tuning fork precision standard several times an hour.

Temperature of each conductor was measured by means of an iron constantan thermocouple inserted into the strands halfway between the measuring points. Temperatures were recorded after each set of three impedance readings. Temperature of the simulated skin section was not measured. D-c resistances were measured with a portable Kelvin double bridge.

Appendix II. Sources of Measurement Error

There are many sources of error in measuring wire impedance by the method described in Appendix I, but it is possible, in most instances, to independently minimize the effects of each error source. These error sources can be divided into three broad categories: failure to establish the conditions specified for the test; failure to precisely reproduce the wire current and voltage at the measuring standards; and inaccuracies within the measuring standards and null detection system.

Failure to Establish Specified Conditions

The effect of incorrect conductor spacing or elevation is difficult to analyze in mathematical terms because it is the same type of effect as the object of this study. Measurements were made with a machinist's scale, and it is believed that each test conductor was spaced and elevated with no greater deviation than ±1/16 inch from the specified location.

Test conditions required that readings be taken with current of only one sequence in the test wires. The check for zero-sequence current was described in Appendix I, but the conditions of only positive- or only negative-sequence current were determined by adjusting the line ammeters for equal readings. The error in reading the ammeters is not more than ±4 per cent, and this error will result in unbalanced line currents or the presence of currents of both positive and negative sequence. If the effect of the skin is neglected, then voltage drop along conductor *A*, measured with unbalanced currents, will be

$$V_a = I_a(R_a + jX_a) - jI_bX_{ab} - jI_cX_{ca} \qquad (3)$$

The impedance of conductor *A* under predominantly positive-sequence conditions is

$$Z_{a1}' = V_a/I_a = R_a + jX_a - j(I_b/I_a)X_{ab} - j(I_c/I_a)X_{ca} \qquad (4)$$

Since both positive- and negative-sequence currents are present

$$I_a = I_1 + I_2 \quad I_b = a^2I_1 + aI_2 \quad I_c = aI_1 + a^2I_2 \qquad (5)$$

Evaluating the ratios (I_b/I_a) and (I_c/I_a) from equation 4 by means of equations 5

$$j(I_b/I_a) = [0.866(I_1 - I_2)/(I_1 + I_2)] - j^1/_2 \qquad (6)$$

$$j(I_c/I_a) = [-0.866(I_1 - I_2)/(I_1 + I_2)] - j^1/_2 \qquad (7)$$

Let $T = (I_1 - I_2)/(I_1 + I_2)$ and substitute equations 6 and 7 into 4

$$Z_{a1}' = R_a - 0.866T(X_{ab} - X_{ca}) + j[X_a + (X_{ab} + X_{ca})/2] \qquad (8)$$

From reference 4, the impedance of conductor *A* with only positive-sequence currents flowing is

$$Z_{a1} = R_a - 0.866(X_{ab} - X_{ca}) + j[X_a + (X_{ab} + X_{ca})/2] \qquad (9)$$

The error due to the presence of negative-sequence current is therefore the difference between equations 9 and 8

$$Z_{a1} - Z_{a1}' = 0.866(X_{ab} - X_{ca})(T - 1) \qquad (10)$$

The quantity T was defined as $(I_1 - I_2)/(I_1 + I_2)$ and by using equations 5 T can be written as

$$T = j\sqrt{3}(I_b - I_c)/(3I_a) \qquad (11)$$

Because the actual magnitude of conductor current does not influence the measurements, it can be assumed that one of the conductor currents I_a is correct, and that ammeter error appears as a deviation in the other two currents I_b and I_c. These two currents can then be expressed as the product of I_a and a complex operator

$$I_b = I_aK_1(\cos\theta + j\sin\theta) \qquad (12)$$
$$I_c = I_aK_2(\cos\phi + j\sin\phi)$$

In equations 12 quantities K_1 and K_2 are the relative magnitudes of I_b and I_c and are restricted by the ammeter error to the range 0.96 to 1.04. Angular quantities θ and ϕ, the phase displacements of I_b and I_c respectively, are not independent because with the neutral return open the sum of the conductor currents must be zero. Using equations 12 this fact can be expressed as

$$1 + K_1(\cos\theta + j\sin\theta) + K_2(\cos\phi + j\sin\phi) = 0 \qquad (13)$$

Rewriting equation 11 in terms of 12 gives

$$T = j(\sqrt{3}/3)[K_1\cos\theta - K_2\cos\phi + j(K_1\sin\theta - K_2\sin\phi)] \qquad (14)$$

Eliminating the angular quantities θ and ϕ between equations 13 and 14 results in

$$T = (\sqrt{3}/3) \times \sqrt{[(K_1 + K_2)^2 - 1][1 - (K_1 - K_2)^2]} + j(\sqrt{3}/3)(K_2^2 - K_1^2) \qquad (15)$$

Substituting T from equation 15 into 10 gives the error directly in terms of the magnitudes of the conductor currents. The error is a complex number and the real part will affect the resistance measurement while the imaginary part affects the reactance measurement.

$$\text{Resistance error} = 0.5(X_{ab} - X_{ca}) \times \left[\sqrt{[(K_1+K_2)^2-1][1-(K_1-K_2)^2]} - \sqrt{3}\right] \quad (16)$$

$$\text{Reactance error} = -0.5(X_{ab} - X_{ca})(K_2^2 - K_1^2) \quad (17)$$

Resistance error attains a maximum value of $0.046(X_{ab}-X_{ca})$ for $I_b = I_c$, and maximum reactance error is $0.080(X_{ab}-X_{ca})$ occurring for the maximum difference between I_b and I_c. The quantity $(X_{ab}-X_{ca})$ can become very significant, being equal to more than the average resistance of some of the unsymmetrical groups which were tested. This discloses that current unbalance is a potential source of serious error; however, if the same current unbalance is maintained during both the positive-sequence and negative-sequence measurements, as is likely, then the effect of the error is eliminated from the average. This can be seen by writing the equivalent of equation 8 for negative-sequence current conditions

$$Z_{a2}' = R_a + 0.866T(X_{ab} - X_{ca}) + j[X_a + (X_{ab} + X_{ca})/2] \quad (18)$$

It is evident that averaging equations 8 and 18 will eliminate the terms containing the quantity T. Because of this, it is reasonable to believe that ammeter error may seriously affect the measurement of conductor impedance for currents of one sequence, but will have much less effect on the average values obtained.

Failure to maintain the supply frequency at 400 cycles does not affect reactance measurements because the measuring circuit actually measures the inductance of the wire, and this is not dependent upon frequency for small variations from 400 cycles. Resistance measurements, as shown by equation 9, do contain a frequency sensitive term $(X_{ab}-X_{ca})$. It was necessary, therefore, to monitor the frequency on a continuous basis and to hold it within 1 cycle of 400 cycles by means of a precision tuning fork.

Failure to Reproduce Conductor Current and Voltage

The measuring standards and null detection system are only capable of comparing the current and voltage which are reproduced at the location of these devices. The current in the measuring standards will not represent the conductor current if the current transformers have any ratio or phase-angle error. Precision current transformers, such as were used in these measurements, have negligible ratio error and less than 10 minutes' phase-angle error within their operating range. Reference 6 analyzes the effect of phase-angle error, and it is there shown that no measurable impedance error results if the phase-angle error is less than 10 minutes.

The primary reason for failing to reproduce the potential drop properly along the test wire is the presence of induced voltages in the loop formed by the potential measuring leads. The most obvious sources of such induced voltages are the test wires themselves, and the only way to minimize the effect is to route the potential measuring leads to a great distance from the test wires before closing the loop. Separation of the conductors and flow of current in the skin aggravate this problem. Measurements were attempted, with different loops, each closed at some known distance from the conductors. There was no measurable difference in readings obtained from loops closed at 12 feet and at 28 feet from the conductors. Accordingly, the shorter loop was selected because it would be less subject to stray fields from outside sources.

Inaccuracies of Standards and Null Detection

The range of inductances encountered in these tests required that the Brooks mutual inductometer be calibrated from 0 to 100 microhenrys. To facilitate accurate reading of this device a pointer 1 foot long was rigidly attached to the movable element, and an expanded scale was carefully prepared. This made it possible to easily distinguish inductance differences as little as 1/4 microhenry. The standard resistor R in Fig. 9, and associated voltage dividers R' and P', were combined into one unit and calibrated from 0 to 0.322 ohm. The scale was arranged to permit an observer to distinguish resistance differences as low as 0.00002 ohm. These standards were constructed specifically for these tests from components having an accuracy of ±0.1 per cent of rated value.

The amplifier of the oscilloscope used for null detection was kept at full gain throughout the tests. Absolute sensitivity of the detector was not measured; however, a comparative estimate was obtained by observing the adjustment of the standards required to produce a noticeable departure from null indication. In no case was this more than 1/4 microhenry on the inductometer. For the least favorable conditions, with small size wire, a change of 0.0002 ohm on the resistance scale was required to alter sensibly the null indication. For larger wire sizes, a change of approximately 0.00005 ohm was always noticeable. In evaluating these sensitivities, it must be remembered that only an 8-foot test length was used and impedances were relatively small. Despite this, it is believed that a substantial gain in measurement accuracy was achieved by measuring along the conductors rather than between them because end connections and bolted joints did not enter into the measurement.

References

1. Impedance Data for 400-Cycle Aircraft Distribution Systems, D. W. Exner, G. H. Singer, Jr. *AIEE Transactions*, vol. 71, pt. II, 1952 (Jan. 1953 section), pp. 410–19.

2. Impedance of 400-Cycle Three-Phase Power Circuits on Large Aircraft and Its Application to Fault-Current Calculations, C. K. Chappuis, L. M. Olmstead. *AIEE Transactions*, vol. 63, 1944, pp. 1213–19.

3. Circuit Analysis of A-C Power Systems, Vol. I (book), Edith Clarke. John Wiley and Sons, Inc., New York, N. Y., 1943.

4. Three-Phase Measurements of Resistance, L. W. Matsch, N. C. Basu, G. R. Horcher. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 350–52.

5. Impedances of Three-Phase, 400 Cycle Aircraft Circuits, 208 Volts Between Conductors. *General Electric Data Folder No. 63004*, General Electric Company, Schenectady, N. Y., 1943.

6. Effective Resistance and Inductance of 3-Conductor Shipboard Power Cables, Samuel D. Summers. *AIEE Transactions*, vol. 67, pt. II, 1948, pp. 1345–50.

7. Wire, Electrical, 600 Volt, Copper, Aircraft. *Military Specification MIL-W-5086*, Nov. 30, 1950.

8. Wire, Electrical, 600 Volt, Aluminum, Aircraft. *Military Specification MIL-W-7072*, Nov. 30, 1950.

Discussion

Bert V. Hoard and **D. W. Exner** (Boeing Airplane Company, Seattle, Wash.): This experimental determination of 400-cycle distribution system impedance including the effects of airplane skin proximity has added considerable valuable information for use of the aircraft industry. This information confirms previously published information on zero-sequence impedances. It indicates larger changes in positive-sequence resistance and reactance than had been previously generally recognized, as the spacing between conductors, their average elevation above the skin of the plane, and the conductor configuration are changed. This added information is of especial value. For example, conductor *AN-0*, the largest size considered, showed an increase in effective resistance as high as 50 per cent when located close to the skin with flat 2-inch spacing as compared to 5-inch elevation and similar spacing. While the effective positive-sequence resistance increased, the reactance decreased by 15 per cent.

In practice, normal designs are such that these extremes usually do not occur. More important, it is possible to avoid or minimize them if it is known under what conditions they can occur. However positive-sequence resistance increases of up to 15 per cent appear possible in practical designs for the larger types of conductor, where flat spacing close to the skin is used and where used increased heating of these circuits should be considered in the design. It should be mentioned that the use of multiple-channel main feeder and distribution systems reduces the degree of this skin effect to a reasonable figure for systems of present-day sizes.

The author states in the section entitled "Positive-Sequence Impedance" that individual phase impedances (for non-symmetrical configuration) may depart rather seriously from this average figure (of positive-sequence impedance). It would have been interesting if there had been included some representative unbalanced impedance information for each of the three wires, at least for the larger size conductors and extremes of elevation and spacing. If this information were included, it is believed that a better understanding of the unbalance factors which are discussed in the paper would be obtained by the reader.

It is of interest to observe that the measured d-c resistance of the *AN* conductor samples tested varied from about 7 to 16 per cent lower resistance than the maximum d-c resistance permitted by the specifications, and for this reason they are lower than usual published d-c resistance values.

Has the author examined experimentally the effects on a-c resistance of certain cable strandings which give a modicum of transposition and the effects of interstrand surface resistance, particularly noticeable in aluminum cable? It would appear that these effects might be of some importance in certain cases.

In the B-52 airplane with a 240-kva system we make general use of 1/2-inch flat spacing, in most cases spaced approximately 1/2 inch above ground plane. Extensive use of multiple-channel wiring makes it unnecessary to use much wire larger than *AN-4*, so that the skin effects which are described by the author have not presented a problem.

Eric T. B. Gross (Illinois Institute of Technology, Chicago, Ill.): The problem of eliminating transposition structures on high-voltage lines has recently led to the evaluation of the resultant unbalances and to a study which started a number of years ago.[1,2] The problem investigated by the author is of a somewhat similar nature. The important difference is the replacement of ground by the aircraft skin section as path for the return current. Analytical expressions for the impedances of such return paths have been derived[3] and may be of interest to the author.

REFERENCES

1. UNBALANCE OF UNTRANSPOSED OVERHEAD LINES, Eric T. B. Gross. *Journal*, The Franklin Institute, Philadelphia, Pa., vol. 254, 1952, pp. 487–97.

2. ELECTROMAGNETIC UNBALANCE OF UNTRANSPOSED TRANSMISSION LINES, Eric T. B. Gross, M. Harry Hesse. *AIEE Transactions*, vol. 72, pt. III, Dec. 1953, pp. 1323–36.

3. GROUND CURRENTS, GROUND FAULT AND GROUNDING PROBLEMS (book), F. Ollendorff. Julius Springer, Berlin, Germany, 1928.

Diesel-Electric Locomotive Wheel Slipping; Causes, Effects, and Methods of Control

R. I. FORT
NONMEMBER AIEE

WHEEL slipping is a serious problem in Diesel-electric locomotives. At best, it results in loss of tractive effort which is the locomotive's only reason for existence. On the other hand, slipping can result in mild to extensive damage to traction motors, rails, and main generators. The Diesel-electric locomotive has an electric transmission capable of converting the constant horsepower output of a Diesel engine into a widely variable tractive effort-speed characteristic. This characteristic (within design limitations and ignoring relatively minor variables such as generator, motor and gear losses) is, theoretically, an equilateral hyperbola with the axes as asymptotes. There are practical limitations at both ends of the curve as to how much of the curve may be utilized, and both limits are important to this discussion. At the slow-speed high-tractive effort end, limits of rail adhesion are encountered, while at the high-speed low-tractive effort end the traction motors have an unfortunate tendency to come apart.

The characteristic of high-tractive effort at slow speed fits into railway train operation very well and is highly desirable. It permits starting long, heavy trains without taking slack or lunging, which practices are not good for car draft gears or contents, whether human or inanimate.

Cars with conventional friction journal bearings, as are the majority of freight cars today, have a relatively high breakaway resistance. Once the car is in motion, this journal resistance drops to a low value and then climbs gradually as speed increases. The combination of these characteristics means that a Diesel locomotive will start a heavy train with ease although going deeply into the traction motor short-time rating. After starting, the train gains speed, quickly at first, and soon gets out of the short-time zone. With improvements in traction motor insulation and cooling, the motor continuous rating is nearing a figure of 25-per-cent rail adhesion. Railroad operating departments are taking advantage of this improvement in hauling-capacity which means that, on hills particularly, locomotives are operated for considerable periods of time very close to the slipping point.

Per-cent rail adhesion is synonymous with coefficient of friction and is an expression in per cent of the possible tractive effort divided by weight on rail of the driving wheels. There are many factors which affect adhesion such as rail condition (new, broken-in, or badly worn—these vary wheel contact area), contaminants (oil, water, ice, leaves, insects, etc.), burned spots on rail from previous slips, weight transfer between axles in trucks and between trucks caused by the tractive-effort couple, binding of moving parts in trucks causing unequal weight distribution, etc. These all have effect at slow speeds. At high speeds, those mentioned are all still effective although in somewhat different order of importance, with the following additions: weight shifts caused by body rocking, spring reaction, and wheel reaction caused by track inequalities.

Wheel slipping falls naturally into two main categories: the starting or slow-speed slip, and the slip at high speed. Although closely allied, they can be considered separately as will be developed. The slow-speed slip results in immediate reduction in tractive effort and may even cause stalling of the train. Doubling hills is not efficient as there is extra mileage and delay to the train involved, and possibly delay to other trains. Any wheel slip must be detected and prompt action taken or the motor may reach destructive speeds. This can happen very quickly. There is a record of a locomotive traveling 6 miles per hour with a motor geared for an equivalent top safe speed of 62 miles per hour going into a wheel slip. In a matter of seconds the axle involved reached an equivalent speed of over 90 miles per hour. That was 50-per-cent overspeed on an 1,800-pound armature with a 16-inch-diameter commutator and 19 inches diameter at the band wire. Centrifugal forces on the armature parts are terrific at such speeds.

Railroads are aware of the seriousness of wheel slipping and recent intensive investigation has brought to light many interesting facts. One of the troubles which set off an exhaustive study on one railroad was burned rail. Routine statis-

---

Paper 54-350, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted July 15, 1954; made available for printing July 26, 1954.

R. I. FORT is with the Illinois Central Railroad, Chicago, Ill.

tical analysis of items of expense indicated an increase in damaged rail caused by engine burns. A closer check on typical districts developed percentages of rail burned ranging from 15 per cent to 68 per cent for an average of 23 per cent of 610 miles of track inspected. To replace the damaged rail would require 33,000 tons of rail and the 1953 allocation for those sections totalled only 22,000 tons.

At first there was some doubt as to whether or not the Diesel-electric locomotive was the sole offender as to rail burn, for steam locomotives also can and do slip. There are certain factors that differ on the two types of power. The larger diameter drivers on steam locomotives provide a larger area of contact with the rail and the mechanical coupling of all drivers helps to stabilize operation at high-tractive effort. On the other hand, the flow of power to steam locomotive drivers is not even as there are four peaks of torque per revolution of the wheels. At present, because of smaller driving wheels, individually powered axles, and high-tractive effort, it is felt that the Diesel-electric locomotive is accentuating this particular problem.

A rail burn may vary from the size of a 25-cent piece in area and a few thousandths of an inch in depth to 3 or 4 inches long and the width of the rail head for deeper burns. These burns affect riding comfort and, more important, are a point at which wheels continually pound, causing a change in rail-head grain structure which can lead to transverse fracture which is a very serious matter.

Theoretical consideration, confirmed by tests, has shown that most rail burns occur at standstill or up to about 1/4 mile per hour. At or above this speed the grinding action of the slipping wheel is spread over a greater area and local effect is minimized. At a speed of 1 mile per hour, the locomotive will move a little over 17 inches in 1 second. Table I, obtained by means of an oscillograph, illustrates the speed with which damage can occur.

The problem is, first, to detect the wheel slip and, second, to do something about it. Automatic means of detection must be provided as events happen too fast for human correction, and the operator is further handicapped by other factors, particularly in multiple-unit locomotives. On a 4-unit locomotive, not uncommon these days, the rear pair of drivers may be 200 feet behind the engineer. Wheel-slip detection devices fall into two main categories: those which measure relative speeds of driving axles through electrical characteristics of their associated motors, and those which measure speeds of driving axles directly through apparatus whose prime purpose is this function.

Table I. Speed With Which Damages Occur

Time of Slip, Seconds	Slipping Speed, Miles per Hour	Depth of Rail Burn, Inches	
		Right Rail	Left Rail
5.10	12.8	0.042	0.034
6.11	12.6	0.027	0.033
2.83	8.5	0.006	0.006

The first group compares voltages across pairs of motors in series, or combinations of even numbers of motors all in series. The comparison is commonly made by a relay comparing the mid-point of motor combinations with a fixed mid-point established by a resistance network across the entire group of motors. This arrangement serves very well for series combinations of motors as it is simple, readily adjustable, uses dependable relays, and can be made as sensitive as practical limitation will allow.

It should be noted that there are certain inherent limitations to the "traction motor" type of wheel-slip detection device. The factors of major importance are wheel-diameter variation and differences in electrical characteristics of the individual traction motor circuits. Wheel diameters may vary considerably, for instance, a new 36-inch wheel is condemned at 33 inches. It is possible, therefore, to be comparing motors connected with wheels nearly 3 inches different in diameter. At 5 miles per hour, this means 51 rpm for the 33-inch wheel, and 47 rpm for the 36-inch wheel, or a difference of 4 rpm. However, at 80 miles per hour the figures become 815 rpm for the 33-inch wheel and 747 rpm for the 36-inch wheel, a difference of 68 rpm. Sensitivity of the relay remains constant on a percentage basis so that, from an absolute standpoint, the arrangement loses sensitivity at higher speeds. A bias, which is proportional to generator voltage and thereby approximately proportional to train speed, can be applied to the relay to increase sensitivity at the higher speeds.

This discussion tends to be academic for, due to difficulties in levelling trucks, etc., most roads hold the difference in any truck to 1 inch on wheel diameter. On the other hand, variation in wheel diameter between trucks is not so important and recent circuits may compare motors of different trucks. The question of difference in wheel diameter had been a moot point on one railroad for some time after the installation of their first streamliner. This locomotive had four motors, the two on each truck permanently connected in parallel. One day it was found that the train would operate with a 36-inch and a $33^{1}/_{2}$-inch pair of wheels in the same truck. Identical meters were installed in the motors of this truck and a round trip was made observing conditions. Electrically, nothing unexpected was found; the predicted current ratio between motors followed wheel diameter ratio closely until the locomotive went into field shunt. Then the ratio reversed, but not seriously enough to issue special instructions. It was not considered necessary to readjust the motor field-shunting resistors as the imbalance was not disturbing.

The second group of wheel-slip detection devices measures axle rpm directly through a mechanical connection to the axle. They may take the form of permanent-magnet generators either with commutators to give d-c output or dry-plate rectifiers to convert a-c output. Another form uses a commutator or cam-operated switches to give pulses of current which may be measured and used to compare relative speeds of two or more axles. This latter type detects differences in rpm so that relative sensitivity increases with speed.

When the wheel-slip detection device operates, power to the spinning wheels must be reduced so that the wheels may regain their grip on the rail. A true wheel slip will almost never correct itself. The kinetic energy stored in the rotating parts is no help and molten metal at the point of contact acts as a lubricant. Prompt reduction of applied power is essential, as acceleration of the slip is very rapid and an early correction limits the attainable speed, thereby minimizing or preventing damage. Most arrangements reduce or cut off excitation to the generator feeding the motor involved and some schemes also reduce the engine speed. At the same time an audible or visual warning is given to the engineer. Voltage-decay time of the main generator is an important factor as it may run from about 0.4 second up to 1.3 seconds. Restoration of power, after the slip has been stopped, must be done with care otherwise a new spin will result. This is accomplished by regaining field excitation and engine speed more deliberately than it was reduced. The engineer is instructed to reduce throttle if repeated slips are encountered.

For a number of years, wheel-slip detection was confined to the series, or slow speed, phase of locomotive operation. The theory was that therein lies the high-torque operation, and at the high-speed, low-tractive effort end of the curve there was no need for protection. Operating experience showed that this portion of the speed-tractive effort curve did need attention so, as the motors were now usually in parallel, a means of detecting imbalances was tried. This was to connect a current-sensitive relay from a point between the armature and field of one motor to the same point on another motor in parallel. It worked, but, under certain abnormal conditions such as grounds in the right places, it would result in full generator voltage on a relay with low inherent resistance and literally blast the relay out of the control cabinet. Then the current type of relay, which compares currents through pairs of motors arranged in parallel, was tried. This, so far, has done well and has been applied to many units now in service.

Wheel slips at high train speed do little, if any, damage to rail but can result in disaster to the motor if allowed to reach excessive speeds. Loss of tractive effort, unusual wear on wheels, and motor flashovers are among the lesser evils. This type of wheel slip can continue for almost unbelievable distances. There are records of wheels traveling an extra 60 miles in a distance of 147 miles. Wheel-slip relays are now being used to initiate automatic sanding. The relay is set to a high degree of sensitivity and tests at starting indicate that it can correct a slip within 1/2 revolution of the driving wheel. This idea is now undergoing extensive service tests. There is no simple solution to the problem but the railroads are joined with the manufacturers in refining and perfecting wheel-slip control.

◆

Discussion

William A. Kirsch (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The author has admirably summarized the current knowledge of causes and effects of wheel slipping and present methods of its control.

Regarding the methods of control, the scheme of operation of one device now under development may be of interest. The functioning of this device is independent of the electrical characteristics of traction motors and therefore avoids the inherent limitations outlined in the paper. The device under development makes use of permanent magnet alternators, one on each axle, driven at axle speed. The output of each alternator is separately rectified and filtered. The resulting d-c signal is proportional to axle speed. The d-c signals from all axles are fed to a voltage-comparing circuit which makes use of germanium diodes, a vacuum-tube amplifier circuit, and conventional relay. A slip indication will be obtained if the difference in speed between any two axles of the unit exceeds 5 miles per hour from any cause whether motoring, braking, or wheel locking. When a locomotive of several units is used, the wheel slip circuits are trainlined so that the speeds of all axles in the locomotive are compared, allowing wheel slip to be detected even if all the axles on one unit should slip simultaneously at the same speed.

The device has been on test for 9 months on one locomotive unit. While testing has not been completed, results so far have been successful. A second test installation will be made in the near future.

R. I. Fort: Information on the device undergoing development mentioned by Mr. Kirsch is very interesting, particularly the fact that comparison among all axles, even with multiple units, becomes practicable. This should reduce to an infinitesimal figure the possibility of simultaneous slips not being corrected.

The Maximum Response Ratio of Linear Systems

PAUL E. PFEIFFER
ASSOCIATE MEMBER AIEE

EXCEPT in cases where pure undamped resonances occur, stable linear transmission or filter systems are characterized by bounded response to every bounded input. What is the maximum response amplitude that can be obtained in a linear, time-invariant system subjected to inputs having limited amplitudes? This question may be stated in another way. To what level must the input to a system be limited to insure that the output (or signal level at any given point in the system) does not exceed a given level? This question is of importance in many physical problems.

The servomechanisms problem described later in the paper prompted the investigation which resulted in the two answers reported in this paper. One of these is a mathematically complete answer with a simple physical interpretation. This answer is in terms of the response of the system to a step input, which can often be determined experimentally with relative ease. While this complete answer is conceptually and physically simple, it is not always easily computed, even when the transfer function of the system is known. A partial answer in the form of an inequality which requires somewhat simpler computations has been found. Although the answer thus obtained is pessimistic, consideration of some numerical examples indicates its probable usefulness.

The Maximum Response Ratio

Suppose b is the bound on the input and h_0 is the maximum instantaneous output amplitude. In mathematical symbols, if $f(t)$ represents the input or forcing function, then

$$|f(t)| \leq b \qquad (1)$$

Similarly, if $h(t)$ represents the output or other observed quantity, then h_0 is the smallest number so that

$$|h(t)| \leq h_0 \qquad (2)$$

The maximum response ratio r is defined to be the ratio h_0/b. For a linear system, r is the maximum output which can be obtained from an input limited to unit maximum amplitude.

With this definition, the question proposed in the foregoing is simply: What is the value of r for the system? Let $w(t)$

Paper 54-371, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted January 4, 1954; made available for printing August 12, 1954.

Paul E. Pfeiffer is with the Rice Institute, Houston, Texas.

This paper is based on a memorandum written for the Bendix Aviation Corporation Research Laboratories. The author is indebted to R. W. Capron and E. C. Johnson for helpful discussions during the investigation and to A. C. Hall for permission to publish the results.

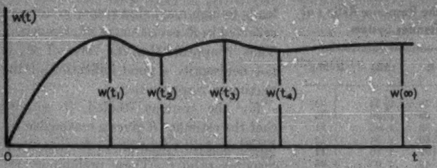


Fig. 1. Response to a unit step function

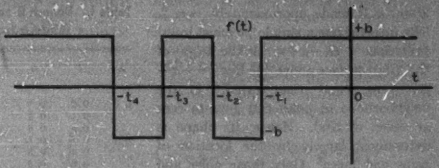


Fig. 2. Function producing maximum output at time t=0

denote the response of the system under consideration to a unit step function input applied at time $t=0$. In general, the step response $w(t)$ will have alternate maxima and minima occurring at times $t_1, t_2, t_3, \ldots$ For example, consider the typical step response shown in Fig. 1. In this case, there are relative maxima at $t=t_1$ and $t=t_3$. Discernible minima occur at $t=t_2$ and $t=t_4$. Since the last of the maxima and minima is one of the latter, the function rises to its final value denoted by $w(\infty)$. The maximum response ratio r for a system having this step function response is given by the expression

$$r=2[w(t_1)+w(t_3)]-2[w(t_2)+w(t_4)]+w(\infty) \quad (3)$$

The maximum response ratio r is equal to twice the sum of the relative maxima minus twice the sum of the relative minima plus or minus (plus in this case) the final value. The plus sign is used before $w(\infty)$ if $w(t)$ increases to its final value and the minus sign is used if $w(t)$ decreases to its final value.

This is the general rule. Although it is stated in terms of a finite number of maxima and minima, this is no real restriction. Even in the case of a theoretically infinite number of oscillations, the contribution after a few terms is usually negligible. For experimental accuracy only those terms which are detectable can influence the results significantly. If the last of these is a relative minimum, then the function is assumed to rise to its final value. If it is a relative maximum, then the function is assumed to decrease to its final value.

Some insight into the significance of the rule may be obtained by considering a signal which will give the maximum response. For the example under discussion such a signal $f(t)$ is shown in Fig. 2. This signal results in the occurrence of the maximum output at time $t=0$. A little study will show that the function $f(t)$ of Fig. 2 may be resolved into the sum of four step functions and a constant. Two of these, occurring at $t=-t_1$ and $t=-t_3$, have an amplitude $+2b$. The other two, occurring at $t=-t_2$ and at $t=-t_4$, have an amplitude $-2b$. The constant term $+b$ may be considered as a step of that amplitude occurring in the remote (theoretically infinite) past. This breakdown leads to a simple interpretation of the result stated in equation 3. The response at $t=0$ is the algebraic sum of the responses to the step components. Those steps which are positive occur at times such that one of the maxima is added at $t=0$. Those steps which are negative occur at times such that minima of the step response are subtracted at $t=0$. The net result is a maximum signal buildup so that the total response $h(t)$ has its maximum value occurring at $t=0$. This picture, which is intuitively sound, may be verified on strictly mathematical grounds. In fact, it was discovered from the mathematics, details of which are found in Appendix I.

A Bound on the Maximum Response Ratio

The exact computation of the maximum response ratio r is tedious, even for a simple system, if oscillations are present in the step response. To alleviate this situation, an inequality has been developed which gives a bound on r. In general, this is not a least upper ground, so that the result is pessimistic. However, as numerical examples show, the bound seems to be reasonable for typical situations that are of interest.

Linear, time invariant systems are characterized by a transfer function denoted by $G(s)$. Here s is the complex variable $s=\sigma+j\omega$. Under the conditions of stability assumed, $G(s)$ is an analytic function for all s to the right of some line $\sigma=\sigma_0$, where σ_0 is negative. The mathematical development of Appendix II shows that the following inequality holds

$$r^2 \leq \frac{1}{2a}\frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} G(-a-z)G(-a+z)dz = R^2(a) = I(a)/2a \quad (4)$$

where

$a<|\sigma|/2$
$-a\leq c\leq a$

It is apparent that the bound depends upon a, which in turn depends upon the real part of the singularities nearest the axis of the imaginaries. Equation 4 shows that it is generally desirable to make a as large as possible; therefore it is of advantage to know σ_0, the real part of the singularity nearest the real frequency (i.e., the ω) axis. This may generally be done with the aid of Routh's criterion or other suitable means, without determining the locations of all singularities of $G(s)$.

The expression for the integral can be written in simpler form by making the substitution

$$G_x(s)=G(s+x) \quad (5)$$

then

$$I(a)=\frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} G_{-a}(s)G_{-a}(-s)ds \quad (6)$$

In most cases of interest, $G(s)$ and, hence, $G_{-a}(s)$ is a rational function of s with real coefficients. The evaluation of $I(a)$ in such cases is possible by methods developed at the Massachusetts Institute of Technology[1,2] and elsewhere.[3,4] Use of tables or systematic computational schedules such as reported in these references facilitates computations. This will be illustrated with numerical examples later in the paper.

An examination of a table of integrals[3,4] will show, however, that calculations increase in difficulty rapidly as the degree of the denominator of $G(s)$ increases. Since it is generally desired to make quick, easy estimates, the question of approximation of $G(s)$ by a lower order function $G^{\circ}(s)$ naturally arises. Under what conditions can $G(s)$ be simplified and how? Some insight may be gained by a study of the expression for $I(a)$, which may be written

$$I(a)=\frac{1}{2\pi}\int_{-\infty}^{\infty} G(-a-c-j\omega)\times G(-a+c+j\omega)d\omega \quad (7)$$

On setting $c=0$ and using the well-known fact that $G(s^*)=G^*(s^*)$, where the asterisk indicates complex conjugates, the following is obtained

$$I(a)=\frac{1}{\pi}\int_{0}^{\infty} |G(-a+j\omega)|^2 d\omega \quad (8)$$

Here, a is made approximately $-\sigma_0/2$. The integral has a simple physical interpretation. $|G(-a+j\omega)|$ is the height of the $|G(s)|$ surface above the line $s=-a+j\omega$. Now $|G(s)|$ rises rapidly in the neighborhood of poles and goes to zero at zeros of $G(s)$. Hence, its shape is determined largely by the location of zeros and poles. Only those poles and zeros near the line $s=-a+j\omega$ are likely to affect the curve $|G(-a+j\omega)|^2$ very much.

In low-pass systems, most of the contribution to the integral will be for $\omega \leq K\omega_0$, where

ω_0 = effective bandwidth
K = numerical factor

The shape of $|G(-a+j\omega)|$ can generally be visualized from the shape of $|G(j\omega)|$ which may be available experimentally. From experimental curves of $|G(j\omega)|$ it should often be possible to approximate $G(s)$ sufficiently well for purposes of evaluation of $I(a)$. Such an evaluation is likely to neglect, in effect, the presence of poles and zeros far removed from $s=j\omega$ and, hence, from the line $s=-a+j\omega$. Rules could be developed to give approximations of proper precision but it is felt that engineering judgment based on the foregoing ideas should provide suitable answers.

An evaluation of the method must involve one important question. The figure R arrived at is an upper bound on the maximum response ratio r, but it is not necessarily a least upper bound. The question remains: How good is R as an approximation to r? Some typical numerical examples might well give a clue.

In view of the foregoing discussion on approximation of $G(s)$, it would appear that a study of representative third-order unit-numerator systems should indicate the usefulness of the method for application to ordinary control systems. Accordingly, systems having the following third-order transfer function will be studied

$$G(w)=1/(Aw+1)(w^2+2\zeta w+1), \quad w=u+jv \tag{9}$$

The transfer function is put in a nondimensionalized form, hence the change of variable from s to w. The function has poles at $w=-1/A$ and $w=-\zeta \pm j\sqrt{1-\zeta^2}$. For $A\zeta \leq 1$, let $a=\zeta/2$. For $A\zeta \geq 1$, put $a=1/2A$. The results of numerical calculations for several representative values of A and ζ are given in Table I.

The method of computation is as follows. The function $G(w)$ may be written

$$G(w)=1/(Aw^3+bw^2+cw+1) \tag{10}$$

Table I. Bounds R on the Response Ratio r of a Third-Order Transfer System

A	ζ	a	R	M_m	R/M_m
0.0	0.2	0.1	3.59	2.54	1.41
	0.4	0.2	1.89	1.36	1.39
	0.6	0.3	1.38	1.05	1.31
0.1	0.2	0.1	3.60	2.54	1.42
	0.4	0.2	1.92	1.36	1.41
	0.6	0.3	1.42	1.05	1.35
1.0	0.2	0.1	2.83	1.83	1.55
	0.4	0.2	1.70	1.07	1.59
	0.6	0.3	1.47	1.00	1.47
2.0	0.2	0.1	2.00	1.19	1.68
	0.4	0.2	1.38	1.00	1.38
	0.6	0.25	1.33	1.00	1.33

where

$b=2\zeta A+1$
$c=2\zeta+A$

Then

$$G_{-a}(w)=1/(Aw^3+Bw^2+Cw+D) \tag{11}$$

where

$B=b-3aA$
$C=c-2ab+3a^2A$
$D=1-ac+a^2b-a^3A$

Tables in references 3 or 5 show that

$$I=I(\zeta,A,a)=B/2D(BC-AD) \tag{12}$$

so that

$$R^2=B/4aD(BC-AD) \tag{13}$$

The R/M_m ratios of Table I are of interest. The M_m value is naturally expected to influence the maximum response ratio. A square wave of amplitude b (i.e., peak-to-peak amplitude $2b$) has a fundamental component with amplitude $4b/\pi=1.27b$. Thus, if the fundamental frequency is equal to the peak frequency, an output of at least $1.27M_mb$ is to be expected. This assumes that the higher harmonics are effectively removed. It is interesting to note that, for the quadratic case ($A=0$), the bound R is only slightly greater than this value, running 1.4 or less. The presence of the real pole at $w=-1/A$ adds complicating effects, however. For $A=0.1$, the pole is sufficiently remote that its effect is quite secondary. For the cases $A=1$ and $A=2$ the situation is quite different, as the numerical results indicate.

It will be instructive to consider the $\zeta=0.2$ case, which is the most oscillatory and has the greatest peaking of the frequency response because of the quadratic term. For $A=0$ and $A=0.1$, the R/M_m ratio is approximately 1.4. As the value of A is changed from $A=1$ or $A=2$, the effect of moving the pole toward the real frequency axis is to pull the response down at the peak frequency. This reduces M_m faster than it reduces the maximum response ratio. The oscillatory nature of the step response for $\zeta=0.2$ leads to high maximum response ratio, as reflected in R, even though M_m is reduced. Thus, it appears that the value of M_m is not necessarily a good indication of the maximum response ratio.

For the systems studied, it appears that the estimate R gives a reasonable indication of the variations in the maximum response ratio r. These preliminary results indicate the desirability of more detailed study, perhaps with the aid of an analogue computer. Whether the apparent fidelity of the estimate R holds for other radically different systems also needs further investigation.

Concerning Applications

The problem which led to this investigation involved a servomechanism which uses a digital or quantized pickoff on the output. The quantized signal may be considered the true signal plus a "disturbance." The latter is bounded by ±1 quantum or digit. Viewed in this manner, the disturbance signal is effectively fed into the input of the servo. Under the assumption of linearity of the equivalent system, the effect of the disturbance can be no greater than the maximum response ratio times the physical amount represented by 1 quantum of signal. If this effect is sufficiently small, then the effect on positional accuracy should be negligible.

It would seem that a knowledge of the maximum response ratio r could be useful in many other situations. For instance, in quasi-linear systems the linear ranges of various elements in the system put limitations on the allowable input level for linear operation. Determination of the allowable limit on the input might well be accomplished by using linear analogues and establishing the permissible ranges experimentally, using the step response as outlined in this paper.

Although the primary concern in this investigation has been the application to control systems and related dynamic problems, it seems that the point of view and the results developed might well be applied to other types of problems. For instance, in problems of determining circuit-breaker requirements, the experimental techniques might be useful.

Appendix I. Mathematical Basis for the Method of Determining the Maximum Response Ratio

In a linear, time-invariant system, the dynamic behavior is determined by the transfer function $G(s)$ or its inverse transform $g(t)$. The filter function $g(t)$ is the response of the system to a unit impulse. In the time domain, the response $h(t)$ to a driving function $f(t)$ is given by the convolution integral

$$h(t)=\int_0^\infty f(t-x)g(x)dx \tag{14}$$

In the following treatment, it is supposed that the system is absolutely stable in the sense that

$$\int_0^\infty |g(x)|dx<\infty \tag{15}$$

For discrete (i.e., lumped-constant) systems, this simply means that all poles of $G(s)$ lie to the left of the axis of the imaginaries. However, it is not necessary to restrict attention to discrete systems.

If the maximum response ratio is defined as in the foregoing elementary properties of integrals show that

$$r=h_0/b\leq\int_0^\infty |g(x)|dx \tag{16}$$

where h_0 and b are defined as in equations 1 and 2. It will now be shown that r is a true maximum response ratio since $h(t)$ can be made to take on the value h_0. Suppose $f(t)$ is the function such that

$$f(-x)=[\text{sign } g(x)]b \tag{17}$$

The function sign $g(x)$ is a function of x which is $+1$ when $g(x)$ is positive and -1 when $g(x)$ is negative. Then, from equations 14 and 16 it follows that $h(0)=h_0$. In other words, for such a function, $h(t)$ takes on its maximum amplitude at $t=0$. It is worth while to ask what kind of a function is indicated by equation 17. The answer is quite simple. Suppose the function $g(t)$ changes sign at times $t=t_1, t_2, t_3, \ldots$ Then $f(t)$ is a rectangular function having the values $\pm b$, with jumps at $t=-t_1, -t_2, -t_k$, etc. If $g(t)$ changes from positive to negative values at $t=t_k$, then the function $f(t)$ jumps from negative to positive at $t=-t_k$, and similarly for changes in the opposite direction.

Now suppose that the response of the system to a unit step function applied at time $t=0$ be denoted by $w(t)$. Then equation 14 shows that

$$w(t)=\int_0^t g(x)dx \tag{18}$$

It is apparent that the points where $g(t)$ changes sign from positive to negative values correspond to maxima of $w(t)$. The points at which the change is in the opposite direction correspond to minima of $w(t)$.

Consider the case that $g(t)$ changes sign a finite number of times. Suppose it does so at $t=t_1, t_2, \ldots t_n$. Moreover, suppose the usual case that $g(t)$ is positive in the first interval and that $w(t)$ settles to a final positive value $w(\infty)$. Then

$$r=\int_0^\infty |g(x)|dx=\int_0^{t_1} g(x)dx-\int_{t_1}^{t_2} g(x)dx+\ldots+(-1)^n\int_{t_n}^\infty g(x)dx \tag{19}$$

In the following, only the integral signs with the limits of integration will be written. It is understood that the integrand is $g(x)$. Using this abbreviated notation, careful examination will show that equation 19 may be written as

$$r=2\int_0^{t_1}-2\int_0^{t_2}+\ldots+2(-1)^{n-1}\times\int_0^{t_n}+(-1)^n\int_0^\infty \tag{20}$$

This is the same as writing

$$r=2w(t_1)-2w(t_2)+\ldots+2(-1)^{n-1}w(t_n)+(-1)^n w(\infty) \tag{21}$$

Equation 20 may be justified by pairing terms in equation 19 and considering the two possibilities, n even or odd

$$\int_0^{t_1}-\int_{t_1}^{t_2}=2\int_0^{t_1}-\int_0^{t_2}$$

$$\int_{t_2}^{t_3}-\int_{t_3}^{t_4}=-\int_0^{t_2}+2\int_0^{t_3}-\int_0^{t_4}$$

$$\int_{t_{2k}}^{t_{2k+1}}-\int_{t_{2k+1}}^{t_{2k+2}}=-\int_0^{t_{2k}}+2\int_0^{t_{2k+1}}\int_0^{t_{2k+2}}$$

For n even

$$\int_{t_{n-1}}^{t_n}-\int_{t_n}^\infty=-\int_0^{t_{n-1}}+2\int_0^{t_n}-\int_0^\infty$$

For n odd

$$\int_{t_n}^\infty=-\int_0^{t_n}+\int_0^\infty$$

Other cases of initial sign of $g(t)$ and final sign of $w(t)$ may be examined in a similar fashion, giving the result of equation 21. Careful comparison will show that this is the same general rule described in the body of the paper.

Appendix II. A Bound on the Maximum Response Ratio

To obtain the desired inequality, the following device may be used

$$|g(t)|=\epsilon^{-at}\epsilon^{at}|g(t)|,\ a>0 \tag{22}$$

The number a is as yet an undetermined positive quantity. It was established in Appendix I that the maximum response ratio is given by

$$r=\int_0^\infty |g(x)|dx \tag{23}$$

If the integrand of equation 23 is written in the form of equation 22, then the well-known inequality of Schwarz may be used to assert

$$r^2\leq\int_0^\infty \epsilon^{-2at}dt\int_0^\infty \epsilon^{2at}g^2(t)dt \tag{24}$$

The first integral is a well-known definite integral having the value $1/2a$, while the second may be viewed as the Laplace transform of $g^2(t)$ evaluated at $s=-2a$. The number a, of course, must be chosen so that the integral exists. Because the response of linear systems is of exponential type, the condition of equation 23 guarantees that such a positive a exists.

Now a theorem on the Laplace transformation of products[5] makes it possible to write equation 24 in the form

$$r^2\leq\frac{1}{2a}\,\frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} G(-2a-w)G(w)dw=I(a)/2a \tag{25}$$

provided

$$a<|\sigma_0|/2,\ \sigma_0<0, \text{ and } \sigma_0<c<|\sigma_0|-2a \tag{26}$$

Here it is assumed that the nearest singularity of $G(s)$ lies at a distance of $|\sigma|$ to the left of the ω-axis. The expression is put in a more symmetric form by the change of variable $w=s-a$. Thus

$$I(a)=\frac{1}{2\pi j}\int_{c'-j\infty}^{c'+j\infty} G(-a-s)G(-a+s)ds \tag{27}$$

Equation 4 is established. The restrictions of equation 26 become

$$a<|\sigma_0|/2,\ \sigma_0<0, \text{ and } -a\leq c'\leq a \tag{28}$$

It should be noted that $G(-a+s)=G_{-a}(s)$, for a real, obeys the conjugate law

$$G_{-a}(s^*)=[G_{-a}(s)]^* \tag{29}$$

Also, if $G(s)$ is analytic for $\sigma>\sigma$, $G_{-a}(s)$ is analytic for $\sigma>\sigma+a$.

References

1. THEORY OF SERVOMECHANISMS (book), **H. M. James, N. B. Nichols, R. S. Phillips.** McGraw-Hill Book Company, Inc., New York, N. Y., 1947.

2. NONLINEAR SERVOMECHANISMS WITH RANDOM INPUTS, **R. C. Boonton, Jr., M. V. Mathews, W. W. Seifert.** *Report No. 70*, Dynamic Analysis and Control Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., 1953.

3. EVALUATION OF INTEGRALS OF THE FORM $\int|Y(i\omega)|^2d\omega$ WHERE $Y(i\omega)$ IS A RATIONAL FUNCTION OF $i\omega$, **D. J. Ritchie, N. Gottesman.** *Bendix Aviation Corporation*, Detroit, Mich. 1952.

4. A UNIFORM APPROACH TO THE OPTIMUM ADJUSTMENT OF CONTROL LOOPS, **R. C. Oldenbourg, H. Sartorius.** *Transactions*, American Society of Mechanical Engineers, New York, N. Y., Nov. 1954.

5. TRANSIENTS IN LINEAR SYSTEMS (book), **M. R. Gardner, J. L. Barnes.** John Wiley and Sons, Inc., New York, N. Y., 1942, p. 275.

◆

Discussion

Abraham Fuchs (Westinghouse Electric Corporation, Baltimore, Md.): Prof. Pfeiffer has preformed a valuable service in providing a straight-forward method for evaluating the maximum response ratio. As pointed out in the paper, the "figure R arrived at is an upper bound on the maximum response ratio r, but it is not necessarily a least upper bound." It is interesting to compute the exact upper bound for the case of $A=0$, using equation 3.

The response of a quadratic to a unit step input may be written as

$$w(t) = 1 - \frac{\epsilon^{-\zeta\omega t}}{\sqrt{1-\zeta^2}} \{ \sqrt{1-\zeta^2} \times$$

$$\cos(\omega\sqrt{1-\zeta^2})t + \zeta \sin(\omega\sqrt{1-\zeta^2})t] \quad (30)$$

Maxima and minima of the foregoing response occur at

$$t = w\frac{n\pi}{\sqrt{1-\zeta^2}} \quad (31)$$

They are equal to

$$w(t_n) = 1 - \left(\epsilon^{-\frac{n\zeta\pi}{\sqrt{1-\zeta^2}}}\right)(\cos n\pi) \quad (32)$$

Evaluating equation 32 at $r=1, 2, 3, 4,$ and for $\zeta=0.2$, 0.4 and 0.6 gives

$$\begin{aligned}&\zeta=0.2\\ &r=2(1.5272+1.1463)-\\ &\qquad 2(0.9227+0.7227)+1=3.056 \quad (33)\end{aligned}$$

$$\begin{aligned}&\zeta=0.4\\ &r=2(1.2538+1.0164)-\\ &\qquad (0.9958-0.9356)+1=1.6776 \quad (34)\end{aligned}$$

$$\begin{aligned}&\zeta=0.6\\ &r=2(1.0948+1.0009)-\\ &\qquad 2(0.99102-0.99992)+1=1.2095 \quad (35)\end{aligned}$$

At most the inclusion of all the higher terms would raise these values of r to

$\zeta=0.2 \qquad r=3.225$
$\zeta=0.4 \qquad r=1.692$
$\zeta=0.6 \qquad r=1.211$

Thus, there is approximately a 10-percent difference between the exact maximum response ratio, and that given by equation 13.

It would be of value to correlate, if possible R with the susceptibility of the system to oscillations which are caused by a step-type feedback (as in the application mentioned or in potentiometer feedback). One would suspect that systems with high values of R might be more likely to oscillate.

There is some question whether the values of $M_m=1.00$ for $A=2.0$, and $\zeta=0.4$ and 0.6 are of significance. M_m is generally limited in its application to transfer functions with $M_m>1.0$. When $1.0>M$ at the natural frequency of the system, other criteria should be used in describing system performance.

Thomas N. Whitaker (University of Houston, Houston, Texas): Dr. Pfeiffer has brought to this problem his skill as a mathematician and his experience as an engineer. I feel that this union of disciplines has given the servomechanisms design engineer a useful tool. However, a rigorous mathematical demonstration is not likely to be light reading, so some of the significance of this paper might be overlooked by the general reader.

Two applications of this theorem present themselves immediately. Consider the problem of setting the limiting (clipping) level on the input signal to a system whose instantaneous output must be kept below some established amount. This clipping level can be determined by straightforward application of this theorem.

A second application was mentioned by Dr. Pfeiffer. This is the problem of accuracy in a linear servo system when the feedback signal is quantized. The instantaneous error in such a system can be considered as the sum of two errors: the error caused by the dynamics of the system when the quantum is set vanishingly small, and an error caused by the size of the basic quantum. A major problem in the design of servos operating with digital computors in the feedback loop is that of determining the size of the largest quantum permitted by the accuracy considerations.

Another possible application is to determine quantitatively the influence of stiction (nonviscous friction) on the performance of a system. This particular nonlinearity causes erratic performance similar in some respects to the introduction of a step function at that place in the system. If, as is usually the case, an upper bound to the stiction can be set, the maximum response caused by this imperfection can be determined.

I was interested in putting this theorem to use in detail on some problem. Brown and Campbell[1] provide a convenient problem for this use. This system is characterized by a second-order transfer function. Results of the calculations of the maximum response ratio for this system for 10 values of damping ratio are listed in Table II. These calculations were made by using a slide rule and, therefore, are not exact. However, the inaccuracy is caused by rounding off numbers and not by omitting terms in the response. The contributions of all peaks (maxima and minima) of the damped oscillations were considered in the sums and then the rounding off was done. This is a possibility in second-order systems but is not generally possible in more complex systems.

Table II. Maximum Response Ratio for a Second-Order System as Function of the Damping Ratio

Damping Ratio	Maximum Response Ratio	R of the paper	M_m
0.1	6.37	7.08	5.02
0.2	3.22	3.60	2.55
0.3	2.18	2.44	1.75
0.4	1.68	1.89	1.37
0.5	1.39	1.57	1.16
0.6	1.21	1.38	1.04
0.7	1.095	1.27	1.00
0.8	1.03	1.23	
0.9	1.005	1.25	
1.0	1.0	1.41	

Equation 3 may be rewritten and extended to

$$\begin{aligned}r=2\{&[w(t_1)-1]+[w(t_3)-1]+\ldots+\\ &[w(t_m)-1]+[1-w(t_2)]+[1-w(t_4)]+\\ &\ldots+[1-w(t_n)]\}-1\end{aligned}$$

where

subscripts 1, 3, 5, 7, 9, of m are odd
subscripts 2, 4, 6, 8, 10, of n are even

The odd subscripts correspond to conditions of local maxima and the even subscripts correspond to conditions of local minima. Each term in the braces represents the absolute magnitude of the departure from unity of the response. Representing this relationship in series form

$$r=2\sum_{p=1}^{N}(D_p)+1$$

where

D_p = the absolute value of the departure of the p-th local maximum or minimum from unity
N = the total number of maxima and minima combined and may be infinite

The particular example I worked out in detail was for a double-energy storage system. For such systems, calculations are particularly simple. Successive maxima and minima in the damped oscillatory response follow at regular time intervals. The ratio of absolute values of successive departures then is merely the value of exponential decay corresponding to this time interval. The departure terms can be represented as terms in a power series. If the series extends over infinite time, their sum can be calculated with great ease because of the function to which this series converges.

Measurements in the laboratory made on such second-order systems can use regular square-wave input signals of uniform period. Generation of this test signal is, of course, easy.

REFERENCE

1. PRINCIPLES OF SERVOMECHANISMS (book), Gordon S. Brown, Donald P. Campbell. John Wiley and Sons, Inc., New York, N. Y., 1948, p. 50.

Paul E. Pfeiffer: Both Prof. Whitaker and Mr. Fuchs have made valuable initial steps in correcting a deficiency of the present paper by providing numerical results comparing the true maximum response ratio r with the unnamed bound R. Confident use of the latter awaits either theoretical or emprical evaluation of the error involved in its use. The initial results are encouraging.

George A. Biernson evaluated graphically the integral expression 16 for r in an unpublished paper. Because his interest lies in other directions, Biernson does not emphasize the role of r as the maximum response ratio so that its significance is easily missed. His graphical method, which involves rectifying the derivative of $w(t)$, is in some ways more useful for computation than that reported in this paper. However, it does not seem to lead naturally to the physical interpretation given in this paper. One of the values of the present treatment is that it points out the nature of the input function which will produce the maximum possible output. The resulting physical insight would seem to be of both theoretical and practical value.

Mr. Fuchs raises an interesting question concerning correlation of r or R with the tendency of systems to oscillate with quantized feedback. It would seem that the question which may be answered is not "will the system oscillate?" but "how large could limit-cycle oscillation amplitudes become?" One is tempted to suggest that the answer is qr, where q is the magnitude of the step or quantum and r is the maximum response ratio of the closed loop system.

Accurate Control of Relative Speed and Cut in a Continuous Process Line

E. G. ANGER
ASSOCIATE MEMBER AIEE

D. L. PETTIT
ASSOCIATE MEMBER AIEE

Synopsis: Magnetic amplifier controls one d-c generator to drive a group of d-c motors to process wallboard and plasterboard. Co-ordination of belt speeds to a small fraction of 1 per-cent uses differential control synchro. Rotating knife synchronism is held by synchro-metering device.

MOST of us are familiar with the wallboard and plasterboard which is used in the interior construction of modern homes, but few are familiar with the manufacturing process which brings this board into existence. More than 20 process lines can be found in the United States in this rapidly growing industry. This board replaces the wood or wire lath and plaster construction, and represents an improvement because of saving in construction costs, greater strength of construction, and an inherent resistance to fire damage.

Wallboard and plasterboard are similar construction materials. Wallboard is plain material with relieved edges usually 4 feet wide, and in lengths from 4 to 16 feet, the most common being 8, and usually 1/2 inch thick. Plasterboard usually comes in narrower and shorter sheets and is pierced with holes to facilitate the application of a plaster coating. The material consists essentially of two outer layers of heavy paper separated by a substance like plaster consisting of gypsum, fiber, rosin, and other accelerating and dehydrating agents used to produce a proper curing cycle.

The Process

The process starts at the mine; a gypsum rock is hauled to a crusher and reduced in particle size; then it is placed on a conveyor and transported to the main mill. Here it is further reduced by a heat and crushing process to a fine powder. It is then conveyed on a system of metering conveyors where the proper proportions of the various ingredients are mixed at the leading element of the main process line for the wallboard itself.

As the metered material reaches the forming roll station, it is mixed with water to the proper consistency in a mixing head and fed out onto the lower roll of paper, as shown in Fig. 1. It then feeds through the forming rolls where the lower roll of paper is folded up and over the edges, and the upper layer of paper is placed on top. As it passes through these forming rolls the paper and mud combination is reduced to the required thickness, and glue is applied to the edge of the paper to secure the upper and lower parts together. It then feeds out onto the first belt section at a rate of speed which is adjustable on the machine described in this paper from 30 to 100 feet per minute. A plastic tape is fed under either side of the soft board to reduce the edge thickness so that in the construction of a home a piece of tape can be placed over the joints without being noticed. This plastic tape proceeds along the first two belt sections until the board begins to set. As the board passes along the first belt section, it is rather soft and pliable and does not set up to any degree of hardness until it is on the second belt section. On the third belt section it is somewhat harder and passes into the live roll section where it is exposed to air for drying of the underside. From here it passes into the punch and knife section.

The punch section is used to punch the holes in plasterboard only and is disengaged for the wallboard process. This punch punches the holes without interrupting the movement of the board. It is driven with elliptical gears so that the punch may enter and leave the board while the punch head is moving along the line at the same rate as the board.

The knife section consists of a lower blade with knives at 90-degree positions rotating in synchronism with the board. An upper knife consisting of a single blade is arranged with a clutch and triggering mechanism to be brought into action by a chain and dog release to cut the board to any desired length in multiples of 1 foot from 4 feet to 16 feet. An additional lost motion mechanism permits the selection of 1/2-foot intervals. The board then passes on to an accelerating section and is fed out into a reversing table where it changes direction to feed through an oven line where curing takes approximately 2 hours to harden and dry it. Tables at the outlet of the oven feed the board through saws which trim and true the ends so that butt joints may be used in construction work. Stacking and bundling then follow for warehousing and shipment.

This process is housed in a building approximately 1,000 feet long. The line itself extends the length of the building and return. The process line is shown in Fig. 2.

Problem History

Numerous problems plague the engineer in his efforts to control the drive

Fig. 1. Forming rolls

Paper 54-387, recommended by the AIEE Industrial Control Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing August 18, 1954.

E. G. Anger and D. L. Pettit are with the Square D Company, Milwaukee, Wis.

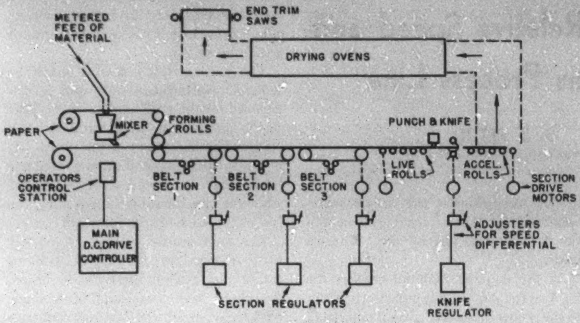

METERED FEED OF MATERIAL
END TRIM SAWS
DRYING OVENS
PAPER
MIXER
FORMING ROLLS
PUNCH & KNIFE
OPERATORS CONTROL STATION
BELT SECTION 1
BELT SECTION 2
BELT SECTION 3
LIVE ROLLS
ACCEL. ROLLS
SECTION DRIVE MOTORS
ADJUSTERS FOR SPEED DIFFERENTIAL
MAIN D.C.DRIVE CONTROLLER
SECTION REGULATORS
KNIFE REGULATOR

Fig. 2. Diagram of process line

for this process. Early machines used a long rope running the length of the machine to power the sections. Speed difference between sections was accomplished by wrapping muslin around the pulleys. Long line shafts were substituted for the rope drive. Slipping and stretching of the rope as well as torsional flexing of the long shaft introduced undesirable irregularities.

Mechanical adjustable speed schemes were employed to vary the speed of the line and the speed of each section. These were subject to speed irregularities because of pulley eccentricities and nonuniformity of the belts used. Also pulleys would wear in the position of the belt for the line speed most frequently used, wearing a groove in the drive pulley at this position. It would then be difficult or impossible to make and hold small changes in speed.

Power synchro systems were introduced to eliminate the long line shaft. These are in current use, in conjunction with mechanical variable speed schemes which have the problems heretofore mentioned. Power synchros, normally operated against field rotation, occasionally start up in reverse. This is especially objectionable in this type of process line.

Line speed must be held within reasonable limits to assure a proper curing cycle. It must be manually adjustable at the forming station to permit the operator to adjust the line speed to accommodate the flow of material out of the mixing head.

As the board passes from the first belt to the second, only a slight amount of tension must be held as the board is very soft and will tear or reduce section if the speed of the second belt is too great. Conversely, if the second belt is too slow, a loop may form, spoiling all the board that passes through the loop. Similar conditions exist between each pair of belts in the sequence but to a lesser extent because the board is increasing in hardness as it sets up along the line. When reaching the live rolls it is quite firm but any excessive slippage may spoil the face of the board.

The most important control problem is that of measuring and cutting the board to length within tolerance as it passes through the knife section. Here the knives must be synchronized with the speed of the line in such a way that they may not only hold the length of cut to close tolerance but they must be adjustable so that the operator may adjust the length of cut. The accelerating section must be adjustable in speed to cause the board to separate as it is cut off by the knife.

The problem was presented to the J. B. Ehrsam and Sons Manufacturing Company wallboard machinery manufacturers, to devise a new approach in an effort to overcome the previously mentioned disadvantages encountered with power synchro systems. The d-c adjustable speed drive system described below was conceived and developed through the co-operative efforts of the Century Electric Company and the Square D Company. This system was installed, and has proved to be a satisfactory solution to the problem.

Main Drive

The selective speed drive includes an a-c main drive motor driving a 50-kw 230-volt d-c generator. The Square D main control is mounted in the motor-generator set enclosure. It provides line voltage starting for the motor-generator set and a 2-stage magnetic amplifier regulator with tachometer feedback for generator field excitation. Current limit is employed for starting. Damping facilitates starting and eliminates hunting. A special cabling plan was used to feed the various d-c motor armatures in order to balance the voltage drop due to the long unequal distances between these motors and the main generator.

Section Co-ordination

Each belt section is driven by its own d-c motor and gear box, each of which receive armature power from the common d-c generator. In order to regulate one belt speed with regard to the next, all belt motor fields are placed in series with

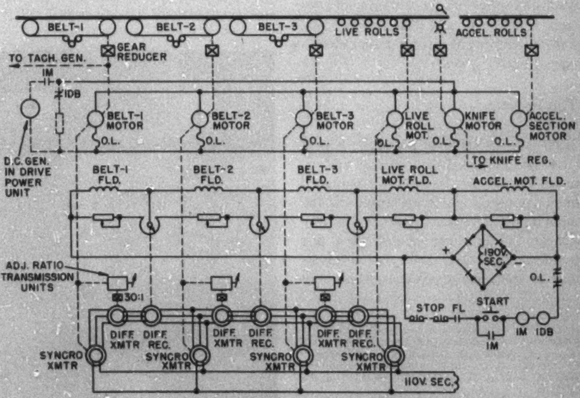

BELT-1
BELT-2
BELT-3
LIVE ROLLS
ACCEL. ROLLS
GEAR REDUCER
TO TACH. GEN.
IM
IDB
BELT-1 MOTOR
BELT-2 MOTOR
BELT-3 MOTOR
LIVE ROLL MOT.
KNIFE MOTOR
ACCEL. SECTION MOTOR
O.L.
O.L.
O.L.
O.L.
O.L.
O.L.
TO KNIFE REG.
D.C. GEN. IN DRIVE POWER UNIT
BELT-1 FLD.
BELT-2 FLD.
BELT-3 FLD.
LIVE ROLL MOT. FLD.
ACCEL. MOT. FLD.
+
190V. SEC.
−
O.L.
ADJ. RATIO TRANSMISSION UNITS
30:1
DIFF. XMTR
DIFF. REC.
DIFF. XMTR
DIFF. REC.
DIFF. XMTR
DIFF. REC.
SYNCRO XMTR
SYNCRO XMTR
SYNCRO XMTR
SYNCRO XMTR
STOP
FL
START
IM
IM
IDB
110V. SEC.

Fig. 3. Diagram of belt regulators

Fig. 4. Potentiometer drive synchro

bridging resistors across each field to obtain speed control by field weakening. The speed relationship of one section compared to the next is controlled by use of a differential control system to vary the resistance in the field shunting circuits.

To insure the proper control of tension, a slip of 1 to 2 inches was specified over the travel of each 125-foot section. Using $1^1/_2$ inches ±1/2 inch as an expression of this requirement, speed matching to ±1 part in 3,000 is indicated. Conventional tachometer speed comparison was not deemed adequate. Instead, by the use of control synchro transmitters coupled to the motors, a position sensing and regulating system was obtained, insuring an inherent zero error in average speed except for that contributed by the adjusting device. Momentary speed errors are minimized by providing a sufficiently rapid correcting action in the regulating controls.

Consider belt sections 1 and 2 in Fig. 3. A synchro transmitter is connected directly to the shaft of each motor. Also connected to the belt-1 motor is a small precision-type adjustable transmission which drives a speed difference synchro differential generator. This speed difference synchro generator is connected between the two speed synchro generators, also in series with a differential synchro which operates as a motor to drive a small potentiometer. Fig. 4 shows the potentiometer synchro drive. The potentiometer is arranged for 360-degree rotation to avoid an excessive current condition in the synchro system should the potentiometer be stopped at the end of its travel.

In observing the operation it will be noted that the speed synchro signal of motor 1 may be modified by the very small differential speed signal correction provided by the adjustable transmission To compare with the speed of synchro generator 1, 2 must run at a speed which equals the speed of motor 1, plus the differential introduced by the speed difference synchro. If this condition is not maintained, the differential synchro motor will rotate and introduce a correction through the potentiometer which is connected across motor fields 1 and 2. It will correct the speeds of the two motors until the differential synchro comes to rest when the speed of motor 2 will equal that of motor 1, plus the speed difference being fed in.

At first it would appear that the response of this system might be somewhat slow, but experience has shown that it is quite adequate considering the requirement of a line of this kind. Initial work on this project was accomplished with a 25-to-1 gear box in the differential synchro drive to the potentiometer to prevent hunting. It was found that this gear box was not necessary, and a direct drive was substituted. Corrections in the system are prompt and without overrun.

It will be observed that the series connection of fields minimizes the effect of field resistance variations with differences in ambient temperatures between motors. Some of these motors operate near the ovens or in open sunlight, whereas others are in a fairly cool portion of the building. With the motor fields in series, the warm-up of any one field affects them all and thereby minimizes the effect, thus reducing the amount of regulating correction required by the control.

Fig. 3 is a simplified schematic of the method of obtaining the speed difference between the belt sections. Actually this system is extended to include belt 3, the live rolls, and the accelerating table. The resistors between the synchro-driven potentiometers permit adjustment to limit the field control range as well as to center the normal operating positions of the potentiometers. An approximate 3-to-1 ratio of field current was found desirable in the adjustment of the system when considering the maximum possible field current of belt 1 compared to the minimum possible field current of the accelerating section. This range of field control was found necessary to regulate adequately at low armature voltages when the current-resistance drop in each armature and its wiring becomes a substantial portion of the total armature voltage. Field strength actually controls torque over a wide range from overhauling on belt 1 to maximum forward torque in one of the later sections in order to

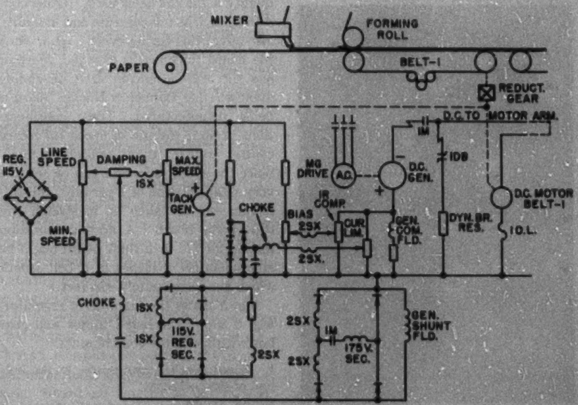


Fig. 5. Diagram of main drive regulator

Fig. 6 (above). No. 1 belt drive

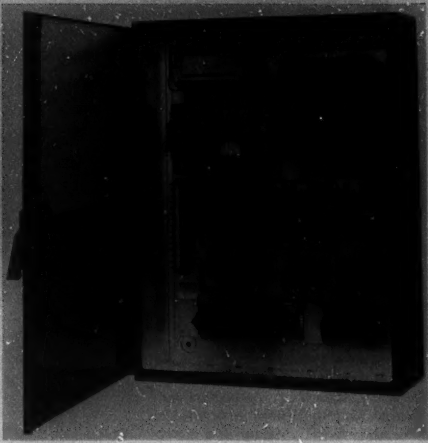


Fig. 8 (right). Knife regulator panel

regulate the required small speed differences.

Tachometer feedback is provided to regulate the belt-1 motor, which controls the generator through a 2-stage magnetic amplifier to produce a constant speed within ±1/2 foot per minute. The system then operates like a master and slave with belt 1 representing the master and the other elements representing the slaves controlled at speed differentials with respect to the master. A speed difference of 3 per cent can be generated at each of the speed difference synchros by suitably setting the ratio of the drive to this unit. Actually, negative differences can be introduced with this system by merely setting the transmissions appropriately, but such settings are not normally required unless belt pulley sizes are considerably out of dimension or the build-up due to tramp material on any one pulley accumulates to the point where it affects the belt speed.

Fig. 7. Synchro drive assembly

While tension control could be applied to a system of this kind, there are certain reasons why speed difference control gives better results. Initially the board is so soft that tension between belts 1 and 2 must be held low enough to cause a regulating problem. There is considerable change in the length of the board as it progresses through its curing cycle so that the board is always swelling or shrinking as it proceeds along the belt line. These positions of swelling and shrinking are not stable points along the line but vary with the speed of the line and the quality of the mix being used. Other conditions such as temperature and humidity also affect this curing cycle. Also, when the board thickness is changed, more or less slip will be encountered depending upon whether thinner or thicker board is being manufactured. These considerations prompt the use of speed difference control rather than tension control. Since the speed difference can be set to a very small fraction of 1 per cent, this difference can be taken up in stretch and slip of the board. A normal condition for slip for 1/2-inch-thick board might be approximately 2 inches in a belt length which is approximately 125 feet.

Fig. 5 shows the main drive regulator. Fig. 6 shows the belt-1 motor and gear box with five drive control units:

1. Speed synchro transmitter in the bottom of the enclosure.
2. Speed difference synchro in the top of the enclosure.
3. Speed indicating tachometer driven from the coupling end of the motor shaft.
4. Line speed recorder tachometer on the other end of the motor shaft.
5. Tachometer generator for line speed control on the other end of the gear box shaft.

Of particular importance in observing the method used is that the regulated difference in speed of the belt sections is a preset small proportion of the over-all line speed. The adjustable ratio transmission has an accuracy of a small fraction of 1 per cent of its maximum setting, and is geared to provide a 3-per-cent maximum correction at that setting. The maximum differential error that the

Fig. 9. Metering wheel and synchro

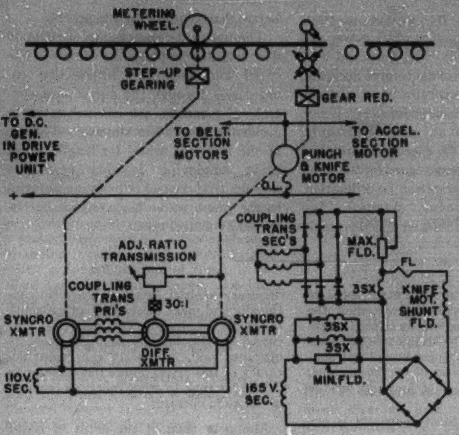


Fig. 10. Diagram of knife regulator

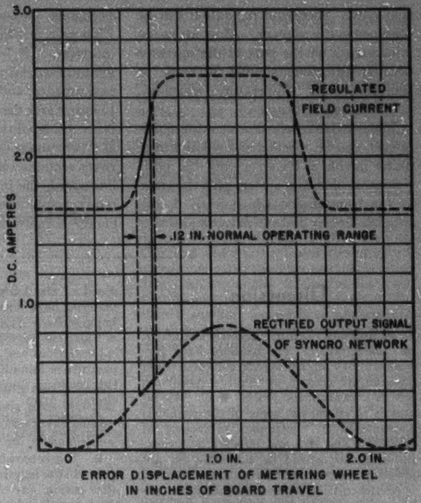


Fig. 11. Knife regulation curves

regulator can permit is therefore less than 0.03 per cent. Since the synchros are driven with sprockets, no further inaccuracy can be introduced because of slip. Instead of employing chains a Gilmer belt drive was used. This is essentially constructed using a steel cable core with a molded rubber conformation for teeth to engage the drive sprockets. The larger sprockets were made of aluminum to minimize the inertia. Taper pin attachment methods were used to prevent slip of sprockets on the shaft; see Fig. 7.

Length of Cut

The knife unit used to cut the board length or to perforate it at the proper intervals is now described. Here the knife motor armature is powered from the same voltage and generator which powers the belt drives. The accuracy requirement on the length of cut was specified at ±1/32 inch or less on board lengths up to 16 feet, allowing an operating variation of less than 1 part in 3,000. This similarly required a position-type regulating system. A magnetic amplifier regulator panel, Fig. 8, controls the field of this knife motor so that the knife is always synchronized to cut the required length. A mechanical triggering device is used to engage the upper single knife so that selected lengths in 1 foot intervals may be obtained at the choice of the operator. Once having chosen the selected length, a metering pulley controls the field through the magnetic amplifier and synchro generator driven by the metering pulley as shown in Fig. 9.

A schematic diagram of the knife regulating system is shown in Fig. 10. This system is similar to the belt regulators in that speed synchro transmitters are employed on the metering wheel and the knife motor with a speed difference being fed in through an adjustable transmission to a speed difference synchro, but a different method is employed to regulate the amplifier. An electric signal from the metering synchro is compared with the signal for the knife motor synchro modified by the speed difference synchro through a set of transformers. The output of these transformers is fed through a rectifier into a magnetic amplifier and forcing circuit to control the field of the knife motor. With this system the knife motor speed can be changed with respect to the line speed by small and accurate amounts through use of the adjustment on the speed difference transmission. In this way the length of board passing through the knife between cuts can be controlled accurately.

Figure 11 shows the knife motor field current plotted against error displacement referred to the null position of the metering wheel synchro. Note that the operating range is a very small steep portion (approximately 0.12 inch) of the total curve. Actually the regulator traverses a considerable portion of this normal regulating range between and during cuts. Board cut lengths do not vary as much as the normal regulating swing because the regulation is a repetitive cycle that is similar from one cut to the next and once adjusted for proper lengths has excellent repeat accuracy.

Also important is the employment of a limited range magnetic amplifier circuit which at the higher current values limits the heating in the motor field and at the lower end limits the approach to a weak field which might cause runaway or instability. Adjustment of maximum and minimum field is obtained through adjustable resistors.

Transient conditions present a problem in regulating the board length. Stroboscopic studies on this installation indicated as much as 1/4 inch of transient movement of the board during the knife cut. Some transient is found when the clutch engages the upper knife for a cut. Eccentricities in the metering wheel drive mechanism may introduce cycle variations into the regulating system. In all, the transients can be tolerated if they subside before the knife reaches the board for a cut. In this particular case it was necessary to rebuild the metering wheel drive mechanism in order to eliminate eccentricities in this drive system. The metering wheel is shown in Fig. 9.

Conclusion

Experiments were conducted to eliminate the speed regulation and work the section co-ordination on torque and tension only. Successful regulation can be obtained in this manner. More attention is necessary on the part of the operator because adjustments must be made to compensate for temperature, warm-up, and thickness of board.

The system described using a speed difference can be considered a position regulator having as its error only the position error of such a system plus the error in the transmission which feeds in the speed difference and which is reduced to a negligible factor by proper gearing. This unit, having only small power requirements, is built with hardened and ground planetary rollers to give accurate performance, and operates in an oil bath to insure long dependable service. When it is considered that any error occurs only as a percentage of the speed difference between sections without regard for the total line speed, the intersection co-ordination requirements are easily obtained. Accuracy of control of relative speed is much better than required by the process.

The use of d-c motors and a common d-c generator simplifies the control of line speed and facilitates intersection co-ordination through the use of field control. D-c motors are not subjected to accidental reversal in starting. The use of the metering wheel for length of cut again employing position regulation through field control reduces inaccuracies to the point where they cannot be observed by a steel tape.

◆

Discussion

G. A. Bumann (Century Electric Company, St. Louis, Mo.): The authors have presented a very interesting and complete paper on accurate control of relative speed and cut in a continuous process line. My intent is to emphasize some of the problems encountered in continuous processing of wallboard and lath and indicate how these problems are solved or simplified by application of a multimotor regulated d-c drive.

At the mixing head immediately preceding the forming rolls, the following ingredients are mixed in metered quantities: Raw gypsum, which is an accelerator; stucco, which is a hydrate of calcium sulphate with a controlled consistency; corn starch, which is used to bond the core to the paper envelope; and paper fibre, rosin foam, and water.

The characteristics of these materials are relatively constant with the exception of the stucco which can vary with the grind, quality of rock, and amount of calcium chloride added, which controls consistency. The paper used in forming the envelope will vary in characteristics since it is subject to a varying storage period under various conditions of humidity and temperature.

Other variables are introduced in an extended line of this nature. For example, hundreds of bearings and various mechanical linkages operating in various temperatures and degree of wear which cause a variation in individual section loading. Since some variables cannot be eliminated we must eliminate any variables which can be controlled so that the machine can be more easily operated. By regulating the speed of the knife to the actual speed of material flow immediately preceding the knife, we have eliminated any error in length of board cut caused by chemical or mechanical variations and changes in ambient temperature or humidity as affects the machine, the chemical reaction of material, or the electrical drive.

Accuracy of length of board cut is important for several reasons:

1. If it is too short, it must be scrapped.
2. If the lengths are trimmed by sawing after curing, any excess length over the desired length, plus an established positive tolerance, constitutes a continual waste, which is considerable over a 120 to 144-hour week at a processing speed of 100 feet per minute. If length is excessive another problem presents itself, namely, the board length might jam at the standard length trimming setup.
3. In the event that the trimming operation is not contemplated, length must be ±1/32 inch or better to allow proper packaging without breakage and ease of installation in construction. For an extended period, board was produced on a machine powered by the drive discussed without trimming. Trimming was eventually introduced only to provide a finer end, since it was determined that the extreme ends were subjected to more heat in the curing cycle and differed in characteristics from the balance of the board.
4. In making lath, the trimming operation is normally by-passed, thus length of lath cut must be accurate.
5. In a similar plant, using a similar machine with a power synchro drive incorporating mechanical speed changes, it was found necessary to check length of board cut continually and make mechanical adjustments to insure accuracy. The drive, as discussed by the authors, simplifies machine operation and, because of the repeat accuracy of this drive, allows the operator to devote his time to other functions required of the knife operator.

Also note that, in the event of board breakage anywhere preceding the knife, the usable material following the break is cut with the same close accuracy as the unbroken board. The same applies to the material at the terminal of a run. Previously, this material was considered scrap.

Maintenance of individual belt speeds and live roll speed, as compared to belt-1, once set is important, as explained by the authors. In drives previously applied, it was necessary to adjust the speed of several sections to compensate for a change in speed introduced in a preceding section. Thus, the operator, upon finding a need for changing a certain section speed, might set up a condition requiring an extended period of manual adjustment. The drive discussed automatically corrects the speed of following sections to maintain speed ratios originally found desirable between those sections, even though preceding section ratios are changed. Hence, this drive provides greater ease of operation and quality control.

Maintenance of line speed once set is important because flow of mix and line speed determine the cubic inches of mix deposited per square inch of paper, hence, thickness and continuity of board. The forming section operator sets mix flow based on line speed. Any serious deviation in line speed causes unnecessary work on the part of the

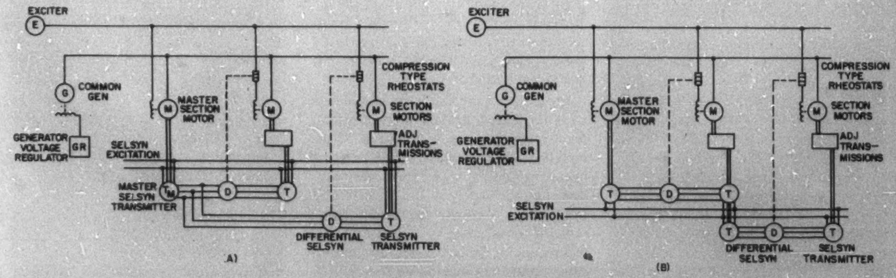


Fig. 12. Typical selsyn type position regulating systems

forming section operator and results in variations in board quality.

In conclusion, by virtue of maintaining accuracy of length of board cut, line speed set, and intersection speed ratios, quality of product is improved and maintained, operating costs are reduced, waste is minimized, and working conditions are improved.

C. D. Beck (General Electric Company, Schenectady, N.Y.): Mr. Anger and Mr. Pettit have presented an interesting paper on a synchro or selsyn regulating system for a plasterboard machine. This application is another example of the present trend which is toward the use of regulated d-c Ward Leonard drives whenever adjustable speed or control of speed, position, tension or torque are required by a machine or process. It is of interest too that the paper and paper converting industry has for many years pioneered the development of many unique d-c drive applications. Today, the pulp and paper industry is second only to the steel industry in installed horsepower of adjustable voltage d-c drives.

The synchro regulating system described is certainly the essence of electrical simplicity and may be entirely adequate for slow speed applications where the load is steady and the reflected load inertia low.

The use of synchros in speed regulating systems for d-c motors on co-ordinated machine drives is not a new development. For approximately 10 years, starting in 1929, the selsyn system shown in Fig. 12(A) was used by the General Electric Company for for sectional paper machine drives. For simplicity, only three sections have been shown, although it was not unusual for a paper machine of that vintage to have eight or ten independently motor-driven sections.

The plasterboard machine drive provides what is often referred to in the paper industry as progressive draw. With this feature, if the speed of one section is increased or decreased, the speed of the following sections automatically changes an equivalent per-unit amount. On at least one paper machine installation this was accomplished by the selsyn arrangement shown in Fig. 12(B). The usual method of obtaining progressive draw was by simultaneous adjustment of all of the following adjustable ratio transmissions from another selsyn system superimposed for that sole purpose upon the system shown in Fig. 12(A).

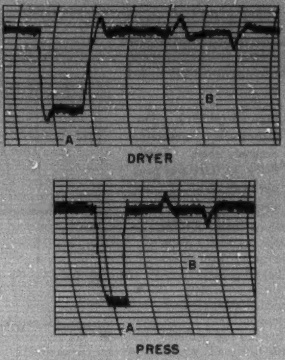


Fig. 13. Response and stability tests on heavy inertia dryer and light inertia press sections. Chart sensitivity = 0.1 per-cent speed change per small division; chart speed = one division in 10 seconds

A—Slack takeup operation
B—Two-to-one change in amplidyne gain to simulate load change exceeding 100 per cent

With the paper machine regulator, Fig. 12(A), the speed of each section motor is compared to a standard or master section. This is usually the preferable system, because it reduces the transient speed differential between sections during power system disturbances, such as in Fig. 13, and when a change is made in over-all machine speed. It also has the advantage that a transient speed error on one section does not further aggrevate the situation by upsetting the speeds of the following sections.

The plasterboard machine system is arranged so that one regulator deliberately disturbs the speeds of adjacent sections when correcting for a speed error. If, e.g., belt-2 motor slows down as a result of an increase in load, commutator film variations, or for any other cause, the regulator attempts to slow down belt-1 motor as well as correct the speed of belt-2 motor. Also, the next regulator in the line is forced into action and, in attempting to help correct the speed of belt-2 motor, upsets the speed of belt-3 motor. In addition, when belt-1 motor was forced to slow down as a result of the initial action by the regulator between the belt-1 and belt-2 sections, the generator regulator further disturbs the entire system by raising the bus voltage. This would appear to be highly undesirable on a machine subject to speed disturbances.

Unless powerful forcing and stabilizing means are provided, stability can be achieved with the synchro system only by slowing down the response of the regulator. This was always the major limitation of selsyn-type paper machine systems because they were essentially ineffective for transient disturbances and often difficult to stabilize if an effort was made to minimize the transient errors by increasing the rate of response. Stability is undoubtedly enhanced in the plasterboard system by a combination of friction in the rheostat and a reduction in torque from the differential receiver, partially as a result of exciting it through the differential transmitter.

As Mr. Anger and Mr. Pettit have pointed out the steady-state accuracy of the system is limited by the accuracy of the adjustable ratio transmissions. Experience on paper machine regulators has proved that this was seldom, if ever, a major limitation. However, the use of the reduction gear and the differential transmitter reduces the errors in this part of the system by a factor of 30 to 1 which in some applications may permit the use of a less expensive transmission.

Because of stabilizing limitations, and the inherently slow response of the selsyn system, it has generally proved inadequate for applications involving higher machine speeds, wide ranges of load inertia, and sudden changes in load.

The newer way is to regulate instantaneous speed rather than position, as with the synchro system. Two commonly used speed systems are shown in Fig. 14. In both, the speed of each section is regulated to a fixed standard. Tachometer generators direct coupled to each section motor are used as a speed-sensing device. An electronic preamplifier and a rotating or electronic exciter are generally used as the power stage of the regulator. Magnetic amplifiers could be used.

The tachometer generator now replaces the adjustable ratio transmission as the

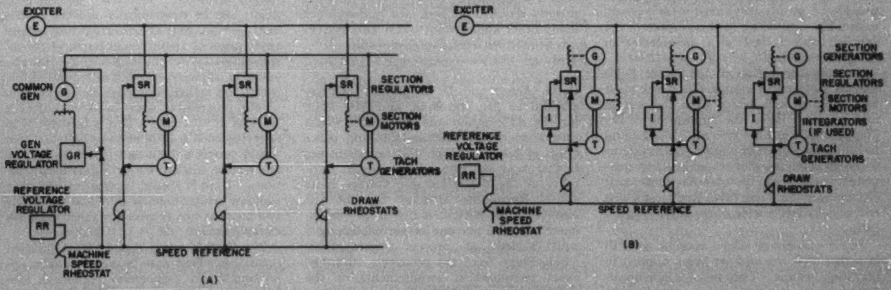


Fig. 14. Typical speed regulating systems

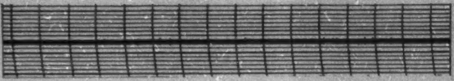


Fig. 15 (above). D-c tachometer generator accuracy chart. Sensitivity = 0.01 per cent per small division; chart speed = one division in 10 minutes

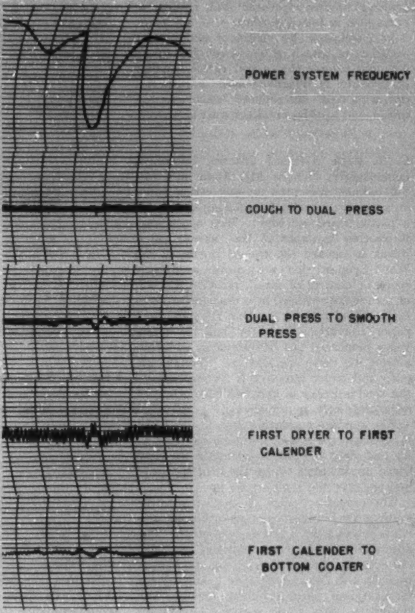


Fig. 16 (right). Speed records resulting from power system frequency disturbance. Chart sensitivity = 0.1 per cent per small division; chart speed = one division in 10 seconds

vulnerable link in the system. However, present-day high-accuracy tachometer generators have an 8-hour drift of less than 0.02 per cent and to date have not proved to be a practical limitation. The transient swings of very short duration, shown in Fig. 15, were caused by speed transient of the synchronous motor system driving the tachommeter generator in the laboratory setup.

On most large machines today, separate generators are used to power each individual section. This system is also highly desirable on small machines where the speed ranges are in excess of 4 or 5 to 1. As the authors have pointed out, with the common generator system the current-resistance drops of the motors and the ineffectiveness of motor field control limit their regulating ability at low speeds. With the separate generator system, each section drive stands alone in space and is completely isolated from any other on the machine. Thus, a disturbance on one section does not effect the speed of any other section.

Speed regulators of the type mentioned with d-c tachometer generators have found extensive use in the paper industry during the last 8 years. At the last count, over 1 year ago, General Electric had 743 electronic-amplidyne speed regulators in service on 107 paper machines alone, with over 20,000,000 section hours of operating experience. By actual measurement, steady-state accuracies in the order of 0.02 per cent were observed on applications where the actual process requirements were proved to be in the order of 0.04 per cent for acceptable machine performance.

The speed system is inherently fast. Its high speed of response together with extremely flexible stabilizing circuits makes it far more suitable than selsyn systems for general application to present-day high-performance machines. The excellent transient characteristics of these regulators are apparent from Figs. 14 and 16 which are charts taken several years ago in the field on actual installations of the common type of generator, shown in Fig. 12(A).

On applications where long time drift and accurate tracking over wide speed ranges are important, an integrating circuit may be added to the preamplifier. The integrator provides a steady-state zero error system, essentially equivalent in this respect to the synchro system, but with no sacrifice in speed of response.

A series connection of the motor fields and the mechanical components which must be located at and driven from the section motor further limits the application of the plasterboard machine system. The general trend today is toward more flexible systems made up of standardized machines and components which can be manufactured most economically on a repetitive basis and which will accomodate a wide range of horsepower ratings.

Power synchro or selsyn systems are still in rather wide use and are being applied to new installations. They provide synchronization during acceleration from standstill to operating speeds as well as during deceleration to standstill, as does the Ward Leonard system with the integrator, shown in Fig. 12(B).

Power selsyn systems usually include a speed regulated Ward Leonard drive for the master section of the machine process. A selsyn generator or synchro transmitter is driven from the master d-c motor and provides motoring and synchronizing torque to the power selsyn receivers or synchro motors.

For slow speed applications where the load is steady and where steady-state accuracy over a narrow speed range is the prime requisite, the author's synchro system will undoubtedly do a creditable job. As they have stated, transient conditions present problems, but they are not severe enough to affect the quality of the product because the accuracy obtained is actually better than the process requires. There are undoubtedly many other applications where this system can be used with equally good results.

E. G. Anger and **D. L. Pettit:** We are grateful to Mr. Bumann and Mr. Beck for the information and comments which they have contributed on this and other types of process drives. These comments further bear out the advantages in flexibility of the d-c adjustable voltage basic drive. Regulating systems can be assembled from devices common to the industry for the purpose of accomplishing a wide variety of process requirements. These range from the relatively slow but exact measurement requirements of the wallboard controller to the high response speed requirements of paper mill drives. In each case the pilot-signalling and amplifying regulator devices become the brain of the system in which any number of d-c motors may be controlled. Pilot devices inherently suited to the process measurement and amplifying regulators with appropriate control range and response can be selected to tailor the system to the process.

Mr. Beck has mentioned a characteristic of the "progressive draw" system of control, in that the advantage to the operator of independence of adjustment of the differential speeds is gained at the expense of additional regulating duty on all the regulators in the system. However, this type of adjustment is of definite advantage on the wallboard machine because of the slow nature of the system disturbances and the limited adjustment ranges involved. In other processes it might not be economically possible or advantageous. It is rather interesting to note that transient disturbances found in the installations did not cause undesirable disturbances in adjacent sections or regulated line speed, but rather permitted close regulation of the progressive draw between adjacent sections at all times. This is contrasted to referring each section back to belt-1, which would permit a disturbance to occur in any one section without correcting adjacent sections. Here it should be remembered that response is intentionally slow and the whole system does not respond to short duration disturbances.

Mr. Beck has also stated that a position regulation effect can be synthetically approximated in tachometer regulating systems by providing devices which furnish the regulating amplifier with a signal which includes the integral of the velocity error. We should, however, like to point out that this cannot always be a substitute for direct measurement of position error in a system where that is the principal requirement and a high accuracy requirement, since the accuracy still depends on the tachometer generator. Mr. Beck cites observed regulation performance of systems regulated by means of tachometer measurements to accuracies as close as 0.02 per cent. However even this figure, when doubled to include the combined errors of two tachometers in a speed comparison system, does not quite meet the design specification of the wallboard controller, which is easily insured by position measurement. The tachometer, too, has a commutator, the variations of which are beyond accurate calculation or compensation. The tachometer generator manufacturer, moreover, has not seen fit to list 0.02 per cent as guaranteed performance. The controller designer and the user of the drive who do not wish to employ devices beyond their accepted range are encouraged instead to use a system which inherently is free of this particular limitation. By using pilot synchros this commutator problem is eliminated. While synchros employ brushes, variation in brush performance affects only slightly the available torque and has no effect on speed comparison.

A fast response regulator with the required accuracy would be prohibitively complex and expensive. The product of this machine is both highly competitive and enjoys a rather limited market. A higher cost cannot be justified for such a control system coping with transient disturbances which have but little effect on the final product.

Our system is based on recognition that refinement of accuracy with ease of adjustment rather than fast response was needed. This is easily accomplished by the use of the position regulation described. Where higher speed of response was required for the knife regulator because of the cutting disturbances, a forcing amplifier extended the system. This system is presented as an improvement with little, if any, increase in cost over the power synchro system as related by Mr. Beck. Accidental reversal, common to power synchros units, is eliminated. Also, the inherent speed irregularities of the associated adjustable pitch pulleys are avoided.

In all systems engineering, prudence demands a compromise to accomplish minimum system complication and minimum initial operating and maintenance costs consistent with accuracy and response characteristics that will have the quality desired in the product of the process line. We believe that this control takes full advantage of these compromises for the best interest of the user and his operating and maintenance staff.

Thermal Developments in Aircraft Generators

CECIL G. MARTIN
MEMBER AIEE

IT has been very adequately and accurately said by many people in the last few years that much more information is required concerning the thermal performance of rotating electric equipment in aircraft. At the same time, a need exists for new devices of improved thermal performance characteristics for many applications. These cover the whole range of aircraft applications from commercial aviation to high-performance missiles. It is the objective of this paper to discuss some of these problems. In particular, the topics are as follows:

1. Rating of air-cooled rotating electric machines.
2. Overtemperature indicators.
3. New developments in thermal design of aircraft generators.

Obviously, all of these cannot be covered in detail so most of the discussion will cover rating with brief coverage of the other two topics.

Rating

The evaluation of the rating problem in air-cooled generating equipment at the Jack & Heintz Company has revolved around a survey of all the temperatures in the machine. This applies particularly to the rotating member of these machines since this has been found to be the predominant critical location of hot-spot temperatures. Some of this work was previously reported in reference 1. These evaluation tests have been carried out on Jack & Heintz model numbers *G29*, *G300*, *G123*, and *G281*. The first of these machines is a 400-ampere d-c generator, the next two are 300-ampere d-c generators differing from one another in design, and the *G281* is a 40-kva *MIL-G-6099* alternator. Each of these machines was thoroughly instrumented with thermocouples in all of the likely hot-spot temperature locations.

The rotating member was instrumented with thermocouples read by means of a rotating thermocouple switch, as shown in Fig. 1. The electromotive force (emf) produced by the thermocouple is transmitted along the hollow shaft of the generator and through a rotating yoke-and-switch assembly which sends the emf to slip rings and brushes of the same basic metals as the thermocouple. The emf is then transmitted to a potentiometer for recording.

The basic environmental conditions which determine rating are inlet-air temperature, air-pressure drop, and rotative speed and altitude. Once values of these variables are selected and set in a given test setup, the load may be increased until safe allowable temperatures are reached. This will determine the maximum safe load for any desired set of environmental conditions. Conversely, if desired, the load and all but one of the other variables may be set, and that one variable may be adjusted up and down until maximum safe temperature limits are reached. The particular technique to be employed depends upon the facility with which different variables may be altered in any given test setup. The basic results will be the same regardless of the specific technique employed.

One of the test setups used is shown in Fig. 2. The generator is mounted on a variable speed drive with the rotating thermocouple switch connected through a rubber coupling to a shaft extension on the generator. The air is supplied through an American Society of Mechanical Engineers orifice installation and an electric heater shown at the top of the figure. The heater is capable of heating and controlling the air to temperatures as high as 250 degrees Fahrenheit (F). The stationary thermocouples are connected directly to a self-balancing potentiometer while the rotating thermo-

---

Paper 54-414, recommended by the AIEE Air Transportation Committee, and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 29, 1954; made available for printing August 23, 1954.

Cecil G. Martin is with Thompson Products, Inc., Cleveland, Ohio.

The quality of the *G123* test results can be attributed to the careful supervision of George S. Stibor, and the commutation limits to Robert J. Shafranck, both of Jack & Heintz, Inc.

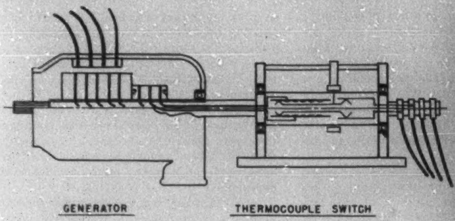


Fig. 1 (above). Test setup schematic

Fig. 2 (right). Safe allowable rating test setup on G29, 400-ampere d-c generator with rotating thermocouple switch

couples are read on the portable potentiometer shown in the lower foreground. The range of variables covered in testing each one of these machines is shown in Table I.

Of prime importance in evaluating the results of the rating tests are two of the machine characteristics. These are the air-flow characteristics and the loss characteristics as affected by rotative speed. For example, the air-flow characteristics of the *G29* 400-ampere generator are shown in Fig. 3. It is noted that, as the speed increases, the flow rate falls off, particularly at the low-pressure drops. The phenomena occurs because of the open riser type of commutator construction which has been traditional in d-c design for some time. These open risers act like a centrifugal fan, opposing the normal direction of flow. This decreasing air-flow rate with increasing speed would result in a decrease in load carrying capacity at the higher speeds even if the losses were to remain the same.

At the same time that the air-flow rate is decreasing, the losses are increasing with increase in speed, as indicated in Fig. 4, for the same 400-ampere generator. It is noted that the losses at 8,000 rpm full load are on the order of 80 per-cent greater than the losses at 3,500 rpm full load. The combination, then, of increasing losses and decreasing air-flow rate with increasing speed on this type of machine at sea level will result in rapid decrease of safe allowable load with increase in speed.

Examples of the results of these rating evaluation tests on the *G29* 400-ampere machine are shown in Figs. 5 and 6. These are repetitions of curves shown in reference 1. Fig. 5 shows the variation in maximum safe allowable load with pressure drop at sea level for 3,500, 6,000, and 8,000 rpm. For comparison, the predictions of commonly accepted rating guides are shown as a dotted line. It is noted, particularly, that the speed effects are very important in determining rating.

Fig. 6 shows the effects of inlet-air temperature on pressure drop required for 3,500, 6,000, and 8,000 rpm full-load operation. Once again, the variations in pressure drop required owing to rotative speed are of a larger order of magnitude, percentagewise, than temperature effects. Thus, rating techniques must consider this basic effect to be reliable.

Detailed repetition here of all the data obtained on all of the machines with all of the independent variables considered would compose a very large volume of work. Fortunately, however, a correlation of all the data on all the machines was established.

It was decided to determine if the rotative speed, altitude density, or any of the other pertinent variables in the machine affected the heat-transfer coefficients appreciable. It has been well established by people working in the field of heat-transfer work that the mass flow of the coolant is the predominant variable in most heat-transfer systems. Therefore, the heat-transfer coefficient should be primarily dependent on the weight flow of air.

This may be evaluated from

Table I. Range of Variables

Machine	Rpm	Pressure Drop, Inches of Water	Inlet Air Temperature, F	Altitude, Feet	Load
G29	3,500 to 8,000	1 to 30	80 to 250	0	0 to 460 amperes
G123	3,500 to 8,000	0.7 to 33	−40 to 105	0 to 60,000	150 to 460
G281	6,000	0.6 to 12.4	−40 to 105	0 to 60,000	20 to 40 kva, 0.75 power factor
G300	3,500 to 8,000	2 to 30	70 to 250	0	90 to 450 amperes

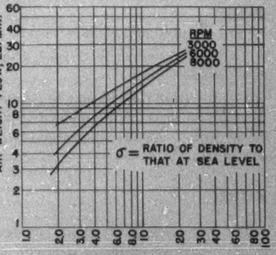


Fig. 3. G29 air-flow curves with radial blast cap

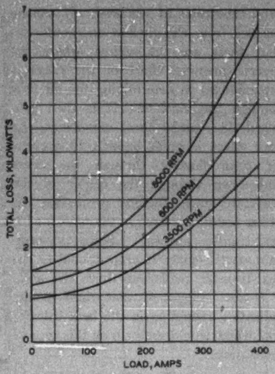


Fig. 4. G29 generator total losses

$Q = HA(T_s - T_a)$

where

Q = machine loss, watts
H = heat-transfer coefficient, watts per square inch, F
A = heat-transfer surface, square inches
T_s = machine temperature, F
T_a = air temperature, F

The product of heat-transfer coefficient times area determined from the foregoing equation, $Q/(T_s - T_a)$, was plotted as a function of air-weight flow on logarithmic paper, as indicated in Fig. 7, for the *G29* generator. The major problem was to determine which surface and air temperatures to use for this plot. It was initially decided that the correlation most likely to succeed would involve rotor-surface temperature, average air temperature, and rotor losses, since the hot spot was located on the rotor. As is noted on the lower left curve, a successful correlation of the data was obtained in this manner. No effects on the heat-transfer coefficient were noticeable other than the air-weight flow. It is emphasized that the full range of data listed in Table I is represented by the data shown. Once a good correlation was obtained with this method it was decided that, since the major portion of the total machine losses (as much as 80 per cent) were in the rotor anyway, perhaps a reasonably accurate correlation could be obtained by using the generator total losses. That this hypothesis is correct is indicated on the upper left of Fig. 7.

While these data were being computed, Prof. Walter Robinson and Dr. E. Naumann suggested that a successful correlation could probably be attained by utilizing the inlet-air temperature as the parameter for air temperature. This suggestion was incorporated and the results are shown in the lower right-hand portion of the Fig. 7. The correlation is undoubtedly successful and, as predicted by Prof. Robinson, resulted only in an increased slope of the basic correlation.

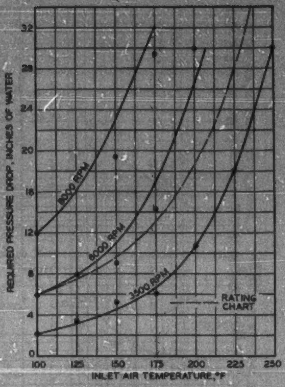


Fig. 6. Pressure drop required for 400-ampere rating of G29 generators with various inlet-air temperatures at sea level

It will be noted that for all three of these correlations, the lower left portion of each correlation indicates a very severe decrease in $Q/(T_s - T_a)$. Examination of the detail data involved in each one of these points has revealed that each of these points occurred at high rotative speed under low-pressure

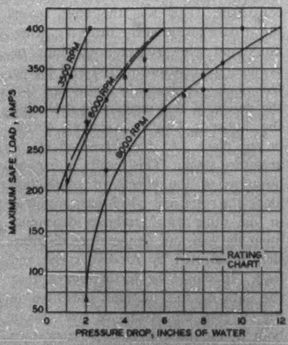


Fig. 5. Safe allowable G29 generator ratings at various pressure drops with 100 F inlet-air temperature at sea level

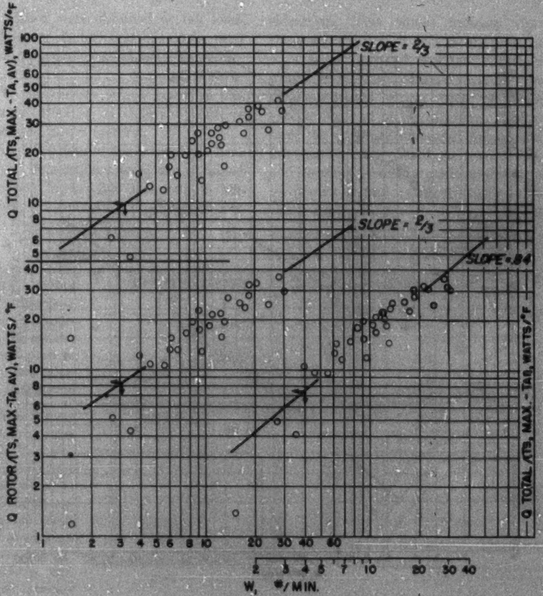


Fig. 7. Heat-transfer correlation, G29 generator

drop conditions at sea level. It will be recalled from the air-flow characteristics that these are the regions where the reverse pressure effect of the risers is predominant in determining the air flow through the machine. It is then proposed that, under these conditions, the riser effect is great enough to prevent air from flowing through the risers and down through the center of the armature stack, thereby precluding the utilization of this area of the machine for heat-transfer purposes. Since the ordinant on the curve is proportional to the heat-transfer coefficient times the area, it is then suggested that these points indicate a reduction in the heat-transfer area which would show up on this type of correlation.

Similar correlations for the *G300* and the *G123* machines are shown in Figs. 8 and 9 respectively, using the same parameters as described for the *G29* machine. It will be noted that the same basic type of correlation has been accomplished. The correlation based on inlet-air temperature has been as successful for these machines as for the *G29* machine. This is particularly encouraging since the inlet-air temperature is one of the basic parameters usually available in the evaluation of a particular installation. The use of average air temperature, on the other hand, involves a trial and error solution of the parameters involved to obtain an allowable maximum loss value. If anything, the use of the inlet-air temperature correlation appears to give better results on all three machines than the use of average air temperature.

This is particularly evident when the data, on the *G123* machine are examined; see Fig. 9. It is noted that the inlet-air temperature correlation is extremely consistent. This was the most recent set of data taken on the three d-c generators and they reflect a constant refinement of instrumentation and testing technique. It should be mentioned that the design of this machine is such that the hot-spot location remains always in one place, i.e., the drive end of the armature stack. This results in a minimization of the effects of conduction along the rotor axis which have occurred in the other two machines depending upon the loss and temperature distribution as affected by rotative speed and load. It should be further noted that the data on this machine presented in Fig. 9 give the results of tests taken both in the Jack & Heintz Laboratory under sea level conditions and in the Wright Field Altitude Chamber under altitude conditions.

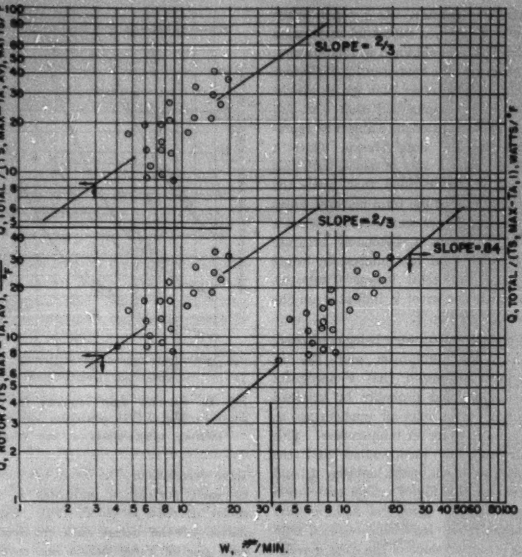


Fig. 8. Heat-transfer correlation, G300 generator

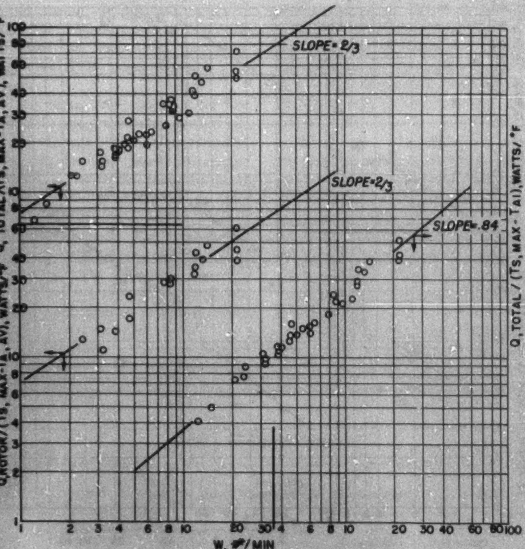


Fig. 9. Heat-transfer correlation, G123 generator

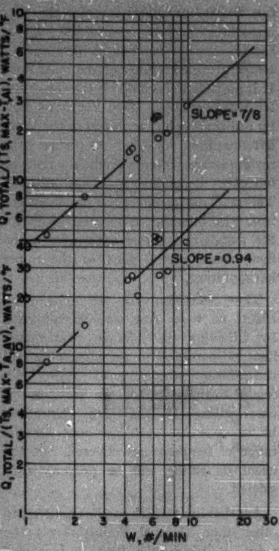


Fig. 10. Heat-transfer correlation, G281 alternator field

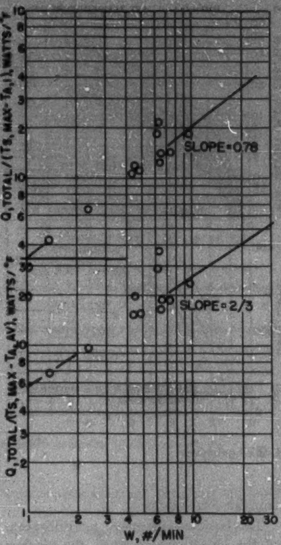


Fig. 11. Heat-transfer correlation, G281 alternator stator

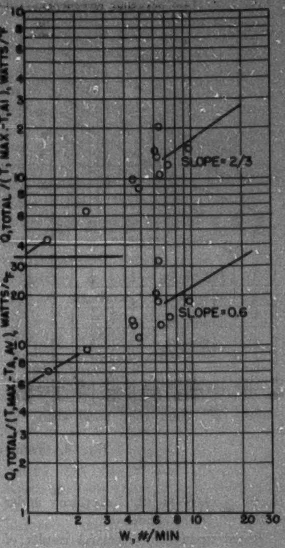


Fig. 12. Heat-transfer correlation, G281 exciter rotor

Although the problem of obtaining successful correlations on d-c generators appears to be relatively simple, as indicated in the preceding figures, the analogous problem on a-c machines is approached with considerably more caution. The loss variation in a-c machines as well as the variation in heat-transfer characteristics is much more susceptible to complication than for d-c machines. In particular, the losses of the a-c machines can vary widely in their characteristics depending upon the speed, the load, and the power factor. These parameters all have different effects upon the losses in the exciter rotor and stator and in the alternator rotor and stator. Thus, there are four basic places where temperature limits can be encountered depending upon the heat-transfer characteristics and the loss characteristics of these particular points. The data which have been obtained to date are too limited to predict the characteristics of these machines with sufficient assurance.

However, Figs. 10, 11, and 12 were prepared for various areas of the machine based on the limited data obtained to check the characteristics. Once again, the ordinates are shown for both average and inlet-air temperatures. It is cautioned that by far the majority of the data shown were obtained at full load, rated power factor, and constant speed for this particular machine. Therefore, the data do not reflect the variations in loss characteristics which occur in the machines because of different rotative speeds and power factors.

Wide variations in the slope of the heat transfer correlation curves are obtained depending upon the location of the maximum temperature point under consideration. This would indicate that the determination of safe rating under a wide range of speed and power factors would necessitate consideration of the detailed losses within each of these specific portions of the machine. Therefore, the data are presented for general information only, based on total losses, and no attempt will be made at this time to predict actual rating techniques or charts until further data are obtained

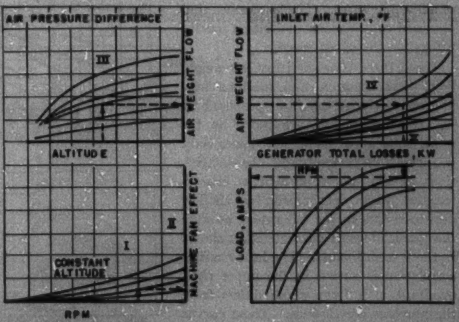


Fig. 13. Rating chart instructions for given revolutions per minute, altitude, pressure drop, and inlet-air temperature

I. From given revolutions per minute and altitude find machine fan effect

II. Subtract machine fan effect from given pressure drop

III. From corrected pressure drop in step II and given altitude find weight flow. IV. From weight flow and inlet-air temperature find total losses. V. From total losses and given revolutions per minute find load

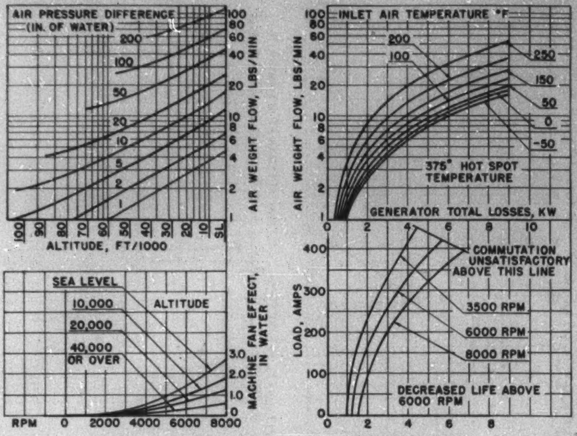


Fig. 14. Rating chart, G29 generator

in the Jack & Heintz Laboratory under wider ranges of conditions.

Based upon the successful results of the heat-transfer correlations on the d-c machines, rating charts along the lines initially proposed by E. O. A. Naumann[3] were attempted.

It will be recalled that the heat-transfer correlations were based primarily upon the losses of the machine, the limiting temperature of the machine, the inlet-air temperature, and the air-weight flow of the machine. It is, therefore, possible to present a curve of air-flow characteristics of the machine which, when the proper variables are considered, will result in a known air-flow rate. This air-flow rate will, for a given limiting temperature and air-inlet temperature, result in a specific value of allowable losses. The loss values thus obtained can be translated into a specific rating for any given speed.

These three characteristics, air-weight flow, heat transfer, and machine loss, approximately in the form proposed by E. O. A. Naumann, are shown in the upper left, upper right, and lower right quadrants respectively on the instructions in Fig. 13. The air-flow characteristics were determined by air-flow measurements in the laboratory as were the loss-load-speed characteristics shown in these co-ordinants. The heat-transfer correlation is shown in the upper right quadrant in a form which differs from the previous correlations. These curves were obtained by calculations using the straight line obtained from the heat-transfer correlations.

The basic requirements for utilization of the charts are to enter the air-flow characteristics curve with a given altitude and pressure difference which determines the approximate weight flow of air. This air-weight flow, along with a given inlet-air temperature, determines the generator losses permissible. The losses, in turn, then determine the load for a given speed of the machine.

This simplified procedure represents the major variables under most conditions but seriously neglects some of the very important secondary variables such as rotative speed effects on air flow and air temperature correction of the mass flow. It is possible to accomplish the correction for rotative speed effect for mass flow by utilizing a plot, as shown in the lower left quadrant of Fig. 13, where the revolutions per minute produce a machine fan effect dependent upon the altitude. This machine fan effect can be subtracted from the pressure differential given for the installation before entering the air-flow characteristic chart in the upper left quadrant.

The correction for air-temperature effects on mass flow was accomplished by adjusting the inlet-air temperature lines proportional to the square root of the absolute temperature ratio with regard to mass flow on the heat-transfer chart. Thus, the mass flow values will be in error by ±10 per cent for the range of temperatures covered, but the effect on rating will be properly considered.

Another limit which can be encountered other than thermal limitations is the point at which the load is increased enough to cause brush arcing or poor commutation. Since the commutation is primarily a function of speed and load, a curve showing this limit can be accomplished in the lower right hand quadrant of the rating chart. This limit has been determined by R. J. Shafranek and is shown in the succeeding rating charts for the three d-c machines investigated; see Figs. 14, 15, and 16.

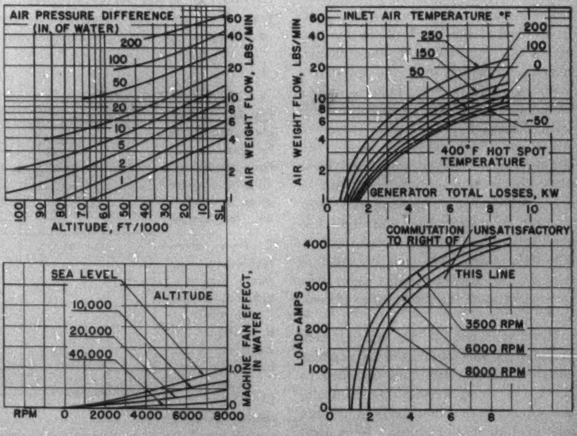


Fig. 15. Rating chart, G123 generator

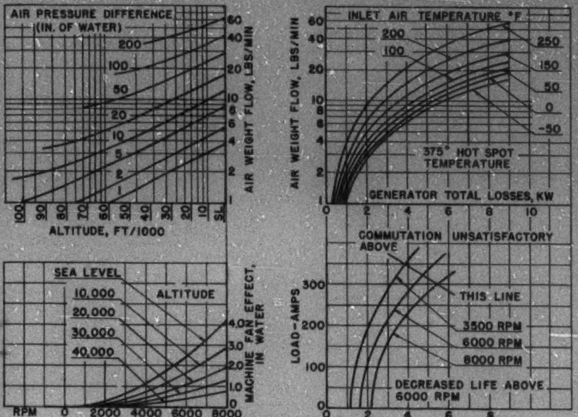


Fig. 16. Rating chart, G300 generator

It will be noted that the rating evaluations thus far obtained take into consideration the limiting temperatures inside the main portions of the machine as well as commutation. Another effect of a thermal nature sometimes encountered is the exit-air temperature. When the exit-air temperature reaches a value greater than is commonly accepted for bearing life of long duration, the bearing performance is endangered unless cool mounting-pad surfaces are provided which keep the bearing temperature within safe limits by conduction. It is possible to show this limit on the heat-transfer correlation portion of the rating chart since the exit-air temperature is entirely a function of the air-weight flow, the generator total losses, and the inlet air temperature. However, instructions for proper interpretation of the curves are somewhat complicated and it is suggested that, for the time being, the exit temperature be computed for the points being evaluated. This can be accomplished by means of the equation

$$Q = WC_p(T\text{ air out} - T\text{ air in})/56.9$$

where

Q = machine total loss, kilowatts
C_p = specific heat for air, 0.24 Btu per pound, F
T = temperature, F

When the exit-air temperature in this equation exceeds approximately 300 F, the air flow must be increased until a value of 300 F or less is obtained. When mounting-pad temperatures in excess of 250 F are encountered, still lower exit-air temperatures must be attained to compensate for this factor. Conversely, when the mounting-pad temperatures are relatively low, the exit-air temperature can be permitted to exceed 300 F. Normally, the complete treatment of this aspect of thermally limited rating is an evaluation peculiar to the particular installation.

When the three d-c generator charts are viewed comparatively, some interesting facts become apparent. It is noted that the machine fan effects are by far the greatest on the *G300* generator and are almost negligible on the *G123* generator. The high machine fan effect of the *G300* generator is caused by relatively small stator air-flow areas compared to the flow areas down through the risers and through the center of the armature. It is this type of configuration which causes the greatest fan effects due to rotative speed. This results in the majority of the air flow going through the machine being influenced by the risers in the direction opposing the normal flow of air. The *G123* machine, on the other hand, sends the air through the machine in parallel paths through the rotor and stator rather than through the riser construction. This is an important advantage when high-speed operation at low-pressure differentials is required of the generator. This can be deduced by going through the rating chart under these conditions to determine comparative ratings.

It will also be observed that the slope of the temperature lines on the heat-transfer correlation is much lower for the *G123* machine than for the *G29* or *G300* machine. This results in a higher allowable loss dissipation for a given weight flow and inlet-air temperature throughout the range of the data indicated. This was accomplished by greatly increasing the heat transfer areas and, thereby, the heat-transfer efficiency of the *G123* machine compared to the other two.

This same characteristic of the *G123* machine having improved rotor heat-transfer characteristics results in another phenomenon of interest. As mentioned previously, the exit-air temperature has an influence on the temperature limited operation of the drive end bearing. For a machine such as the *G123* which is cooled efficiently on the rotor, the limitation will more often be found to exist on the exit-air temperature than for the other type of machine. This type of phenomenon has been common since

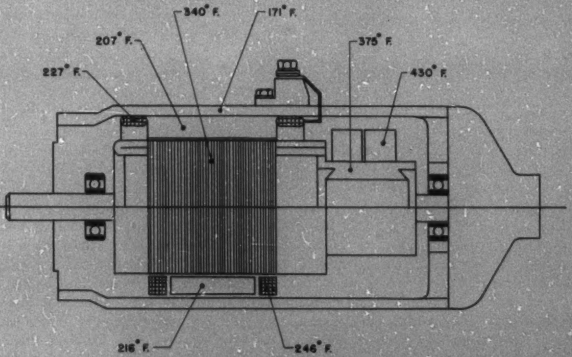


Fig. 17. Example of temperature distribution on a typical d-c generator

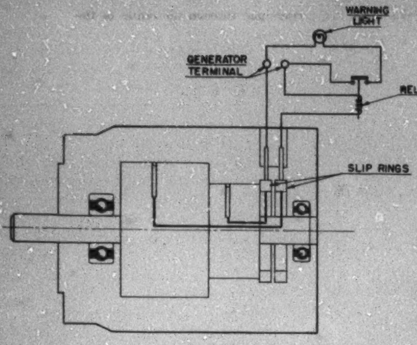


Fig. 18 (left). Generator rotor overtemperature indicator

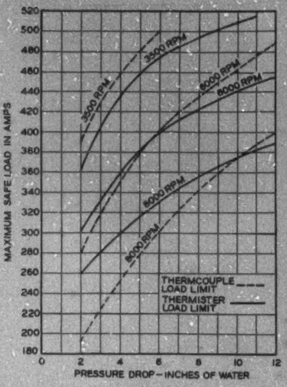


Fig. 19 (right). Overtemperature indicator performance on G29 generator

the early days where, when one part of a design was improved, the emphasis on further improvements was shifted to another part of the design.

It is appropriate to emphasize that the ratings of the *G300* machine were determined on the basis of the known temperature limits of the *G29* machine (375 F). This was reasonable at the time these tests were initiated since the only laboratory machine available for testing had similar insulation, and life-temperature characteristics of the better insulations were not as well defined. Even though the quality of insulation in this series of machines was subsequently improved and the machine qualified to military specifications at higher ratings than those indicated, the time and money already invested in instrumentation on the test machine dictated that it remain unchanged. Therefore, should the data be applied to a *G300* generator, it ought to be recognized that roughly 20-per-cent increase in loads is permissible. The *G29* and *G123* generator performance data are applicable as presented.

Overtemperature Indicators

Although this topic may be a little premature, a correlary to the study of temperatures throughout the machine is the possibility of devising an overtemperature indicator which will reproduce the results of the rating evaluations. A device of this type is being investigated in the Jack & Heintz Laboratory with the promise of successful completion of the problem.

A schematic of a typical d-c generator is shown in Fig. 17, where the pertinent component parts and their operating temperatures at rated load and speed under rated cooling conditions are illustrated. Prior to discussing the over-all temperature distribution, it is pertinent to point out the 430 F brush temperature shown when the commutator temperature is at 375 F. With radial-type blast cap, the air-flow distribution is such that a large percentage of the flow is down through the bottom of the machine opposite the brushes indicated in Fig. 17. For this reason, it has been found at Jack & Heintz that the brush temperatures should be measured at the side of the machine corresponding to the blast-cap inlet. This results in the highest brush temperatures measurable around the periphery of the commutator. In some instances, temperature gradients in the brushes in the order of 100 F have been noticed depending upon the peripheral location of the thermocouples in the brushes.

In addition, brush temperatures varying through wide ranges under fixed and stabilized environmental conditions are often found in testing practice of d–c generators. This is an area which has not been completely defined to the knowledge of the author, but may be attributed at least in part to widely varying brush friction coefficients.

When the load and speed are varied either together or independently, some alteration of the temperature pattern indicated is experienced. Generally, as the rotative speed goes up, the commutator temperature tends to rise faster than the generator armature stack. On the other hand, when high loads at low rotative speeds are encountered, there is a pronounced shifting of the hot-spot temperature to the armature stack. Since under all useful load conditions the hot-spot temperature is in one or the other of these two locations, some means of sensing these temperatures must be devised unless a satisfactory correlation can be found with some stator measurement of temperature. To date, no such correlation has been found. Then, if a requirement is based upon the protection of the armature from overheating, some means must be found to protect both the commutator and the armature stack. Either of these could fail because of overtemperature if only one of them were being protected under some operational conditions. It was decided to try to accomplish this by locating small thermistors in the armature, one in the commutator and one in the hot-spot location in the armature stack.

This configuration is shown schematically in Fig. 18 where the position of the thermistors is indicated with the leads coming out to slip rings and brushes through a relay to generator terminals. This relay can then be utilized to actuate a warning light or a circuit breaker, as required by the particular installation. The thermistors are connected in parallel so that, when either hot-spot location becomes predominant, the resistance will fall off enough to close the relay. Since the resistance decrease of thermistors with increasing temperatures is steep, the temperature of the hot spot is reasonably well protected with a device such as this.

To determine how accurately this basic concept would protect a generator, tests were run with the rotating thermocouple switch on the 400-ampere *G29* generator. The thermistor circuit was connected through this rotating thermocouple circuit and the maximum safe load determined for the thermistor operation. The data thus obtained were compared with the safe ratings determined by the rotating thermocouple switch on a previous run. These data are shown in Fig. 19 where the maximum safe load is plotted as a function of the pressure drop for various speeds. Both the thermocouple and thermistor load limit determinations are indicated. It is observed that, even though this is a greatly enlarged scale, the correlation between the two methods is quite good.

The variation shown at 8,000 rpm and low-pressure drops is attributed to the critical nature of the cooling at this point. That, is, under these conditions when the riser effect is large, small changes in pressure differential occur across the machine which would permit some variation under even fairly closely held test conditions. It is believed that the thermistor is almost as accurate as the rotating thermocouple switch for this type of temperature-sensing operation.

It must be admitted that this device is not yet perfected to the point where it could be sold commercially. Since altitude-treated brushes are necessary for the slip ring assembly to prevent dusting as is encountered on other altitude brushes of a nontreated nature, some excessive filming under sea level conditions is occasionally encountered which has a high resistance to the flow of current. This does not happen very often but does occur frequently enough so that the basic reliability of the system is endangered.

New Developments in Thermal Design of Aircraft Generators

It is believed that the present state of rotating machinery development for aircraft applications can be divided roughly into four categories. These are as follows:

1. Commercial.
2. Present military fighter and bomber.
3. Future military fighter and bomber.
4. Missile applications.

In general, the commercial applications follow along in severity behind the current military fighter and bomber applications. Thus, with relatively minor changes, the commercial needs are met with equipment already developed and proved. Present activities in development of equipment for meeting the needs of present military fighter and bomber requirements will not be discussed to any great length here. These developments consist of a line of high-rotative speed machines thermally improved with regard to insulation, thermal stability, and heat-transfer efficiency. Also available are a fairly complete line of high-altitude inverters and several developments on high-altitude motors and regulating equipment.

For future applications in military fighters and bombers, it is believed that the liquid-cooled systems for rotating machinery will see greater and greater use. Generally speaking, the configuration consists of a piece of rotating electric machinery cooled by a recirculating liquid. This recirculating liquid will most commonly be the engine oil or the hydraulic oil on the airplane. The heat picked up by the oil will then be rejected in the same manner and to the same heat sinks as the engine and hydraulic oil, whichever may be the case. The advantages of this type of operation are as follows:

1. Recirculating liquids are commonly used in aircraft to collect many of the various scattered heat loads generated by engine bearings, gear boxes, constant speed drives, and hydraulic systems and to reject this heat to a suitable heat sink such as ram air, fuel, cabin exhaust air, and the exhaust air from air-turbine-driven accessories. For an engine-driven and mounted accessory like electric generating equipment to follow a set of similar requirements seems a reasonable solution to a long-standing problem.

2. Generators can use, and in fact require,[8] the same working fluids as modern turbine engines, gear boxes, and constant speed drives.

3. The 200- to 300-Btu-per-minute heat load of present 12-kw d-c and 40-kva a-c generators is small compared to the 2000-Btu-per-minute, and higher, recirculating liquid heat loads of present-day turbine engine installations. Expanding these current heat-rejection systems by the additional 10 per cent required appears to be the most efficient and compact means of solving many present installation and environmental problems of electric generating equipment.

4. One of the more significant advantages is realized between the generator and the gear box or engine upon which it is mounted. As long as the gear box or engine utilizes a lubricant compatible with good heat-transfer characteristics, a single cooling circuit is possible, employing the already available pumps, sumps, and external plumbing.

5. The large 3- to 4-inch ducting and still larger shrouds of blast-cooled equipment are eliminated in favor of 1/2-inch tubing.

6. The elimination of the necessity for large internal air passages in the generator enable the diameter and length to be reduced by 10 per cent or more, depending on the size of the machine. This, in combination with item 5, can contribute greatly to relieve the congestion of present engine mounting pads or, more likely, leave them as congested but occupying less aircraft engine frontal area.

7. The momentum drag normally incurred because of loss in energy of the ram air as it passes through the ducts and passages of air-cooled generators can be largely eliminated in a well-designed liquid-to-air heat exchanger system and can be totally eliminated if other heat sinks are utilized.

8. The higher heat-transfer coefficients obtainable with liquids can be often utilized to reduce the size and weight of a thermally limited machine far beyond the 10 per cent mentioned in item 6.

9. In applications requiring the use of an expendable evaporant type of cooling, nonfreezing solutions are often prohibited because of detrimental effects on the machine varnish, bearings, and brushes. Care must also be exercised to prevent the accumulation and freezing of water when it is used. All of these problems are more susceptible to solution when the evaporation takes place at the heat exchanger of a recirculating liquid system.

10. A recirculating liquid system can be adapted to reject heat to any available heat sink and to switch from one to another as required. This is a powerful asset under an extreme range of environments such as encountered by many current aircraft.

11. The present necessity for designing a new air-cooled machine for each new and more severe environmental requirement is eliminated. Only the heat exchanger size and location need be considered.

To illustrate some of these points, a comparison of a 40-kva air- and oil-cooled alternator is shown in Fig. 20. The oil-cooled version is much smaller although the electric portions of the machines are identical. The air-cooled alternator is the Jack & Heintz *G281* and the oil-cooled version is the Jack & Heintz *G190*.

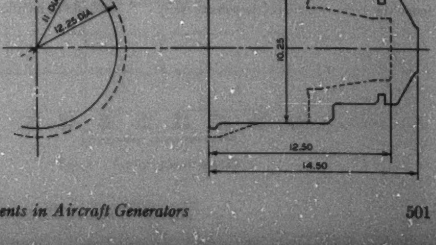


Fig. 20. Envelope comparison of 40-kva air- and oil-cooled alternators

—— Air-cooled machine

- - - Oil-cooled machine

A major disadvantage, of course, is the lack of comparable experience on this type of design in view of the long experience on air-cooled equipment. However, Jack & Heintz has been working on this type of development for more than 4 years, and is rapidly accumulating data and experience on design techniques for several new developments of this type.

For the missile applications, which normally involve extremely high temperatures and high altitudes, the problem of finding a suitable heat sink for rejection of heat becomes more and more severe. This means that, when the flight duration is extremely short, a device of a thermal lag nature may be utilized. An a-c generator of this type has been built with a rating of 8 kva, 120/208 volts weighing 42 pounds, and it will operate on the order of 30 minutes without any cooling at all. Depending on the temperature of the material prior to the firing of the missile, this time can be extended or, conversely, diminished if the temperatures are high at the start of the operation.

Another possibility along missile lines and, in a sense, of the same basic nature as the true thermal lag machine, is the evaporative-cooled machine. Basically, an evaporant machine is a thermal lag device with greater thermal capacity obtained by the use of the latent heat of vaporization of the coolant. The coolant may be made either a part of the basic machine or piped in from a separate source. To date, Jack & Heintz has built and successfully operated two evaporative-cooled designs. These are the *G75* 15-kva a-c generator, and the *G188* 30-kva a-c generator. The *G75* weighs 40 pounds and the *G188* weighs 60 pounds.

References

1. Loss-Temperature-Environmental Relationships for Aircraft Generators, **Cecil G. Martin.** *AIEE Special Publication S-57*, Oct. 1953.

2. Rating Tests of a 40 KVA Alternator, **E. O. A. Naumann.** *WADC Technical Memorandum Report WCLE-53-338*, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

3. 1953 AIEE Aircraft Technical Conference. *AIEE Special Publication S–57*, Oct. 1953, "Survey of Evaporative and Liquid Coolants for Rotating Electrical Machines," **Cecil G. Martin, Valentine F. Hambor,** Paper No. 7.

◆

Discussion

R. A. Yereance (Boeing Airplane Company, Seattle, Wash.): Most of this paper is concerned with d-c generators. I have had little experience with the thermal evaluation of this type of machine, but some comment seems in order.

Among the environmental conditions which determine rating, I believe ambient-air temperature and mounting-pad temperature should be included. Although the effect of ambient temperature on the internal temperatures of a blast-cooled machine is not great, in some modern aircraft, particularly in missiles, the ambient temperatures are so high as to require consideration. Mounting-pad temperature is not listed among the environmental conditions although it is indirectly considered where bearings are discussed. It should be emphasized that lowering the exhaust-air temperatures will not lower the bearing temperatures if the drive mounting pad on which the machine is mounted is at a high temperature.

In our tests we have obtained data which indicate a change in apparent heat transfer with altitude. Although this effect was not noted by the author within the range of his experimental data, as he states in the section on "Rating," he did obtain an apparent change in heat-transfer characteristics under some extreme operating conditions. He states that, under some conditions, portions of the machine become ineffective as heat-transfer areas, which essentially means a change in apparent heat-transfer characteristics. The fact that this effect did not show up in his tests indicates only that, for the machines tested, under the test conditions, the effects were negligible. Nonetheless, the possibility of an apparent heat-transfer coefficient variation occurring is obviously present and apparently, for the machine we tested, was a function of altitude and occurred in the more nearly normal operating range.

Mr. Martin also mentions that, "Wide variations in the slope of the heat-transfer correlation curves are obtained depending upon the location of the maximum temperature point under consideration." He had previously stated, in the same section, that the losses, and locations of the losses, in a-c machines vary widely with load, speed, and power factor. He has therefore offered evidence to support our contention that, under various conditions, there may be an apparent change in heat-transfer coefficient. It seems only reasonable to extend this line of thinking to the conclusion that the apparent heat-transfer coefficient is a function of altitude as both the effective heat-transfer area and the location of the hot spot in the machine may change with altitude because of the increased cooling air velocity necessary to maintain constant mass flow as the altitude is increased. This increased cooling air velocity may change the cooling air distribution within the machine, which is the factor directly responsible for the change in apparent heat-transfer coefficient. It might be noted that other experimenters working on a-c machines similar to the one we tested have also noted this altitude effect.

Instrumenting a machine rotor with thermocouples is probably an excellent method of obtaining rotor temperatures but, as in our case, is often impractical as we do not own the machines we test and are not free to alter them to instrument them in this way.

It should be mentioned that commutation may in some cases be a function of altitude, humidity, etc.

It seems doubtful if the overtemperature warning system discussed in the section entitled "Overtemperature Indicators" will find wide acceptance. It involves the addition of extra slip rings to the machine, which is not easily done on a compact machine, and present rotor temperatures are very close to the maximum operating temperatures for thermistors. The practicability of any warning system, unless it can be made very simply, will be trouble free, and will require no maintenance, seems questionable.

Cecil G. Martin: Mr. Yereance has made an appreciable contribution by his thoughtful analysis and comments. The additional rating parameters of mounting-pad temperature and ambient temperature do have effects on rating which should be considered ultimately. The object of this paper was to explore basic performance parameters.

The phenomenon of redistribution of air and corresponding alteration of heat-transfer coefficient is a complex one, but the explanation and analysis presented by Mr. Yereance appear reasonable, and further detailed experimentation should yield very interesting results.

The instrumentation of machine rotor with thermocouples contributes to the knowledge of the temperature distribution in a rotating part in a manner which cannot be duplicated by resistance techniques. Ultimately sufficient information should be available based on specific temperature measurements to make resistance measurements completely indicative of performance. I concur that any protective system should be extremely reliable and simple.

Environmental Testing of Small Electronic Components at High Sound Intensities

F. MINTZ
NONMEMBER AIEE

M. B. LEVINE
ASSOCIATE MEMBER AIEE

THE use of jet and rocket engines in aircraft and guided missiles creates a significant noise problem. Disturbance caused by noise is not new to the aircraft industry. In general, the major problem has been, and still is, the effects of noise on human beings. Increasingly, however, the effect of noise fields on mechanical systems, in particular electronic equipment, is becoming important.

Fig. 1 shows the sound spectra found aboard and in the proximity of a U.S. Air Force jet aircraft. Curve *A* shows the average sound level in each octave band measured approximately 10 feet from the tailpipe of the jet engine operating at full thrust. Curve *B* is the sound pressure level measured for the same aircraft at the inboard equipment bay. The noise level at the equipment bay approaches 140 decibels (db). These two curves do not represent the worst posssible conditions but may be considered typical of conditions found in service. The sound intensities encountered are sufficiently high to affect the reliable operation of electronic components, particularly vacuum tubes. In view of the amount and complexity of the electronic equipment aboard a modern military aircraft or guided missile, this lowering of reliability leading to malfunction or failure becomes an important problem.

Intelligent mechanical design of electronic components intended for air-borne applications requires data as to the environmental conditions encountered in use and, even more important, data on reliability of existing electronic components and systems under these conditions. The Mechanics Instrumentation and Vibration group at Armour Research Foundation (ARF) is engaged in an experimental program for the U.S. Air Force concerned with this problem.[1] In connection with the acoustic testing phase of this program, a convenient and simple laboratory device has been designed and built for obtaining moderately high-intensity sound fields.

Existing Methods of Obtaining High-Intensity Sound

A number of possible methods may be considered for obtaining high-intensity sound fields for acoustic testing of electronic components.[2–6] Some of these are discussed in the following.

USE OF THE JET ENGINE ITSELF

Although this would provide realistic simulation of the jet engine sound field, for a particular engine this method does not lend itself to economy, flexibility, ease of installation, or ready availability.

USE OF SIRENS AND WHISTLES

Sirens and whistles are frequently utilized for the production of high-intensity sound. They are found in use today for air raid and other warning purposes, and, at various laboratories, for acoustic experimentation. However, the sirens have several drawbacks. To produce a sufficiently high sound level to be practical for component testing, in general, requires a relatively large and costly installation. Secondly, the siren generates basically a line-spectrum sound (rich in harmonic content), although certain research sirens may be conveniently driven and controlled over a wide range of speeds and frequencies. Whistles have similar limitations. The use of these devices thus limits the types of testing that could be accomplished.

AIR-MODULATED SPEAKERS

A recent development in the audio field is the air-modulated loud speaker. This device holds great promise as a high-intensity speaker. At the present time, however, several factors militate against its use. First, it is not readily available commercially. Second, in addition to the usual electronic equipment associated with audio work, the air-modulated speaker requires a compressed air system. Finally, the flow of air associated with this type of generator is incompatible with the design of the chamber as originally, and now, envisaged. The ARF sound chamber embodies a fourth approach, based on the reverberant chamber principle and utilizing inexpensive commercially available components in a rather simple installation.

The ARF Sound Chamber

The basic principle of the ARF sound chamber is as follows: Consider an enclosed space containing a sound generator radiating sound at a given power *P*; the sound intensity *I* will build up until the power dissipated at the walls (and in the air) is equal to the input power. If, as a simple approximation, the intensity is assumed to be uniform throughout the enclosure, its value is given by the expression[7]

$$I=\frac{P}{a}$$

where *a* is the total absorption in units of area. If *P* is given in watts and *I* in watts per square centimeter, the absorption should be expressed in square centimeter units.

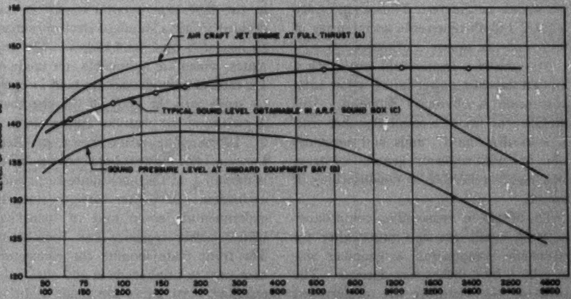


Fig. 1. Typical sound spectra

Paper 54-422, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing September 15, 1954.

F. MINTZ and M. B. LEVINE are with the Armour Research Foundation, Chicago, Ill.

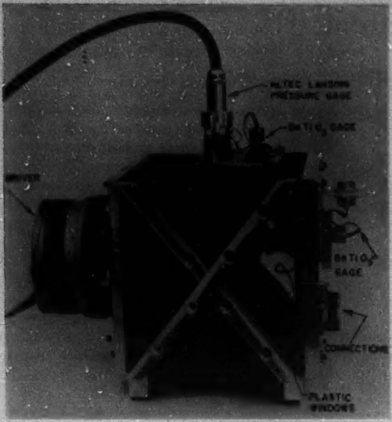
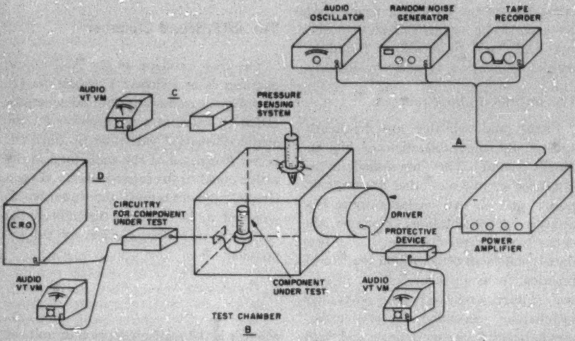


Fig. 2 (left). ARF high-intensity sound chamber

Fig. 3 (below). ARF high-intensity sound chamber and associated equipment



To establish a framework for discussion, it is worth noting that the reference level for air-borne sound, 0 db, is 10^{-16} watt per square centimeter. Thus, an intensity level of 140 db represents an intensity of 0.01 watt per square centimeter.

To achieve a high sound level with a modest acoustic power input, a low total absorption is obviously required. This dictates the use of a small chamber with acoustically "hard" walls and construction, i.e., with walls and general construction highly reflective or nonabsorbing of sound.

To obtain a reasonable compromise between small size and practicality for electronic components, a chamber size of 6 by 6 by 6 inches was selected. The sides of the chamber are constructed from 12-gauge sheet metal and are reinforced with 1/2 by 3/4-inch aluminum ribs. The original design envisaged the construction of the entire chamber from 1/2-inch aluminum plates; for simplicity, in the first experimental model it was decided to use a standard electronic chassis box of 6 by 6 by 6 inches. The end plates, which are removable, are made of 1/2-inch aluminum and are held in place by studs into the body of the chamber. A picture of the chamber appears in Fig. 2. In the configuration shown, an Altec-Lansing *730A* 40-watt loud speaker driver is mounted on the back plate to provide excitation. In a later configuration, an experimental Jensen unit of somewhat higher power capacity has been used. The front plate mounts the connectors necessary to bring leads in and out of the chamber. All joints are made tight in an effort to assure complete closure of the chamber.

Instrumentation directly associated with the chamber consists simply of the pressure measuring system. Since it would be impractical to mount the microphone of a commercial sound-level meter within the chamber, special microphones are used. Three gauges may be seen in Fig. 2; two are Barium Titanate pressure sensors of ARF design,[8,9] the third is an Altec-Lansing condenser microphone designed for measurement of fairly high acoustic pressures. Sound pressure level in the chamber is determined by reading voltage output of the microphone on an audio-frequency vacuum-tube voltmeter (vtvm). Since the chamber was first constructed, windows have been added at the test end. These allow optical measurements to be made on the component under test. The windows are of plexiglass and are sealed tightly to the chamber to prevent losses.

Operation of ARF Sound Chamber

In operation, the ARF sound chamber along with its associated equipment may be considered as four distinct groups of components. These groups, illustrated in Fig. 3, are as follows:

SIGNAL SOURCE AND POWER AMPLIFIER

The basic signal may come from one of three sources: an audio oscillator which provides single-frequency sine-wave input, a random noise generator which provides sound of a general nature covering the frequency range of 20 to 20,000 cycles per second (cps), or a tape recorder which could be used to provide special signal inputs, e.g., recorded sounds of a jet engine. The signal source drives a power amplifier which provides the 40 to 50 watts for electric power required to operate the driver at maximum rated input. Between the power amplifier and the driver a fuse is provided to protect the driver from overload and also a shut load resistor to prevent damage to the amplifier should the driver fail. An audio-frequency vtvm is placed across the driver input for use in ascertaining the input power.

THE SOUND CHAMBER

This group consists simply of the sound box and its driver.

THE PRESSURE MEASURING SYSTEM

This system consists of the Altec-Lansing pressure gauge, its power supply, the two ARF gauges, and an audio-frequency vtvm to read the gauge output. A cathode-ray oscilloscope is used to monitor the wave form of the sound field.

The Component Under Test

The component under test is located in the sound chamber in the area of the windows. It is supported about $1^{1}/_{2}$ inches away from the wall on an aluminum bracket. Since the microphone is located elsewhere in the chamber, the sound pressure acting on the component may be different from that sensed by the microphone, particularly at eigenfrequencies of the chamber. This introduces errors in the result which could be minimized by adjusting the location of the component. The fixed bracket was used, nevertheless, for the sake of simplicity in construction and use. For the case of random noise excitation, the location of the component is not a critical matter as the sound field in the chamber is probably rather uniform.

Leads from the component to its circuitry are brought out through the connectors on the back plate, thus preserving a tight seal on the chamber. The circuitry for the component under test is of necessity located outside the chamber. Instrumentation for detecting the effect of the sound field upon the test component is connected to this external circuitry, the type of instrumentation depending on the component. The oscilloscope and audio-frequency vtvm, shown in Fig. 3, are used with electron tubes.

Calibration of the sound chamber is accomplished by supplying a known electric power at a given frequency to the driver and reading out the sound pressure level in the chamber by means of the pressure gauges. Fig. 4 illustrates the frequency response of the sound chamber at several values of constant power input. The variations in intensity at the higher frequencies are caused by resonances within the chamber. With approximately 50 watts of electric input, sound pressure levels of 145 to 150 db are obtainable. With the random noise generator used as signal source, this same power level yields a sound level of 140 to 143 db in a frequency band of 20 to 20,000 cps.

Operation from a sine-wave source is particularly useful in determining resonances of components. Fig. 5 is based on actual data taken at 2,300 cps, which is one of the chamber resonances. At this frequency, a level of 150 db is obtained with about 7 watts' electric input. For operation at a single-frequency off-chamber resonance the curve shown in Fig. 5 is skewed downward. Measurements of the Q of the chamber at resonance were made by the bandwidth method; Q's of 47 at half-wave resonance, 80 at full-wave resonance, and 111 at the second-harmonic resonance wave found. This is consistent with the sharp peaks seen on the frequency-response curve in Fig. 4.

Referring again to Fig. 1, curve C is an average sound pressure level obtainable in the chamber both with random noise generator and the audio oscillator input. This is well above the sound level encountered at the equipment location.

Results of Tests on Components

Electron Tubes

Type 6J5WGT

Since many electron tubes are microphonic, they were selected as the first components to be tested in the chamber. Tube type *6J5WGT*, a military version of the standard *6J5* triode, was chosen as vibration tests had been made on this tube.

The tube was mounted in a standard octal socket at the window end of the chamber. The tube was tested in operation as a capacity-coupled amplifier with self-bias. The bias, plate voltage, filament voltage, and plate current were selected in accordance with Joint Army-Navy specifications; the Joint Army-Navy specification of 75 millivolts of noise was used as the failure (or malfunction) criterion. The noise voltage was read with an audio-frequency vtvm and the wave form was viewed on an oscilloscope to determine the nature of the noise. Testing was carried out both with the audio oscillator supplying single frequencies and with the random noise generator as input.

Five tubes were tested, utilizing a 120-db sound field and variable-frequency input; then two of the tubes, one exhibiting average microphonic output and the other exhibiting excessive microphonics, were subjected to constant-frequency and random noise excitation at various sound pressure levels. The past history of each tube was known from the time it was received from the supplier. Two of

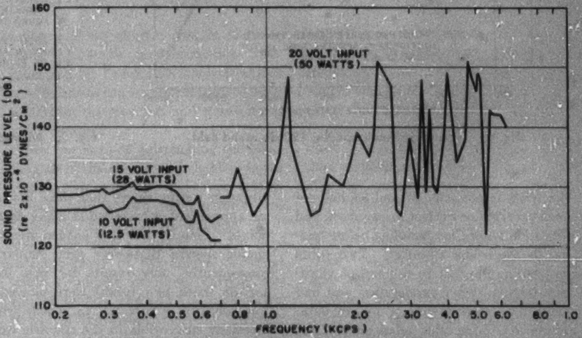


Fig. 4. ARF sound chamber, sound pressure level

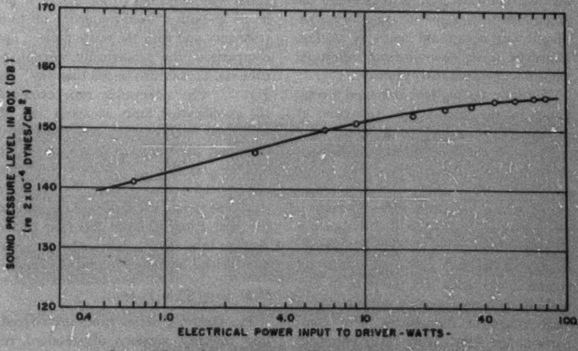


Fig. 5. Typical operation of ARF sound chamber, 2,300 cps

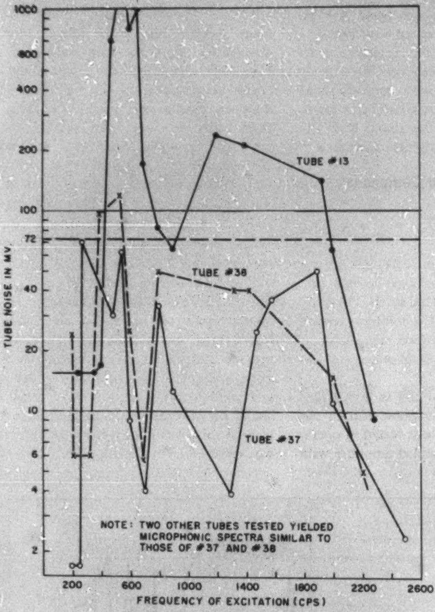


Fig. 6. Tube microphonic spectra, 120-db sound field

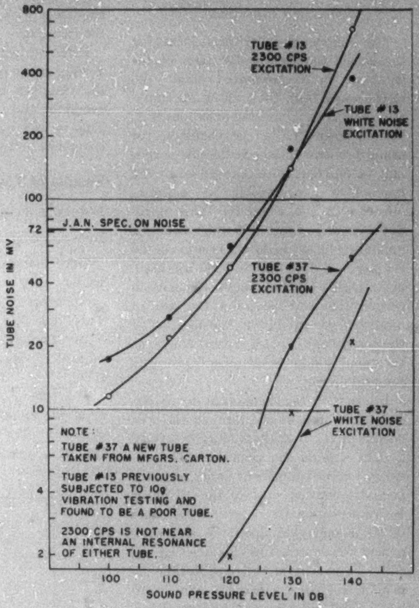


Fig. 7. Microphonic output of two tubes in high-intensity sound fields

the tubes tested, nos. 37 and 38, had not been operated or subject to any shock and vibration other than handling before testing in the sound chamber. Two tubes had been subjected to 8 hours of vibration testing at 5 g acceleration and had yielded no malfunction or failure data. These were acceptable tubes under Joint Army-Navy specifications. The fifth tube, no. 13, had been subjected to 10 g acceleration testing and had exceeded the Joint Army-Navy specification for microphonic output after 4 hours of testing. However, upon removal from vibration, the tube operated satisfactorily.

Vibration testing had indicated a resonance of the grid support structures at about 250 cps in the type-*6J5WGT* tube, with some tubes exhibiting a second resonance at about 500 cycles. The spectra of tube microphonics, as illustrated in Fig. 6, show these resonances along with several others not disclosed in the vibration testing. A level of 120 db is only moderately high in terms of jet engine sound fields, yet the data in Fig. 6 show high microphonic output over a large range of frequencies. For comparison, the Joint Army-Navy specification limit for microphonic output under vibration testing shows that these limits are approached or exceeded at many data points. Four of the tubes yielded microphonic spectra that were very similar; consequently, only results for nos. 37 and 38 (both new tubes) were plotted. The behavior of tube no. 13, which was especially susceptible to the acoustic excitation, is also shown.

A tube considered typical of the average tubes was subjected to increasing sound levels at both a fixed-frequency off-tube resonance and with the white noise. For comparison, the acoustically susceptible tube no. 13 was also tested this way; see Fig. 7. The susceptible tube exceeded the specification limit at about 125 db while the average tube did not exceed this limit until a level of about 145 db had been reached. Excitation with white noise yielded results similar to that of single-frequency excitation in the case of the poor tube, while in the case of the average tube the single-frequency excitation yielded higher microphonic output.

6000 Series Tube Types

The 6000/CT series tubes are ruggedized military versions of standard receiving and transmitting tubes.[10] Three tubes types representing this preferred line of vacuum tubes were subjected to the high-intensity sounds as produced in the ARF sound chamber. As with the *6J5WGT* samples, the tubes were tested operating in the production testing circuits as specified in specification *MIL-E-1B*.[11]

Fig. 8 is a plot of the microphonic output of representative samples of the 6000 series tubes (along with a *6J5WGT* for comparison) when subjected to excitation at a single-frequency off-tube resonance at increasing sound pressure levels. Because of the difference in failure level for each of the types tested, the dimensionless ratio, per cent of failure level is utilized as an ordinate for the curve. This allows direct comparison of the four different tube types. All tubes had a microphonic output exceeding the failure level below 160 db; the 6005/6095 type failed below 120 db.

Fig. 9 is a plot of the microphonic output of the same tube types when subjected to white noise excitation at increasing sound pressure levels. Again failure or malfunction occurred at moderate sound levels. Microphonic spectra taken for the 6000 series tubes yield

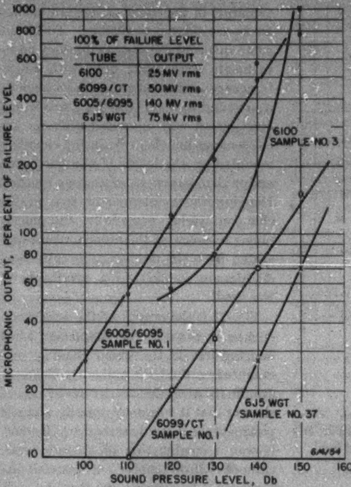
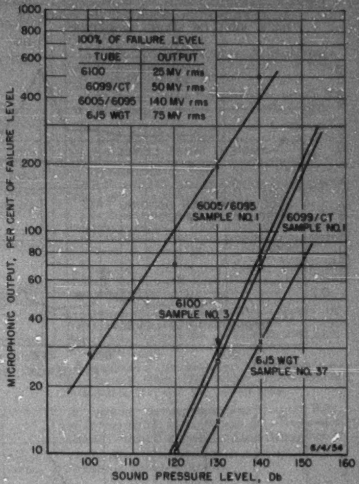


Fig. 8 (left). Microphonic output of various tubes in high-intensity sound fields, 2,300 cps excitation

Fig. 9 (right). Microphonic output of various tubes in high-intensity sound fields, white noise excitation



curves similar to those for the *6J5WGT* tube; see Fig. 6. Resonant frequencies common to a given type are easily apparent.

These data are the results of preliminary testing of tubes There are not enough results yet to allow drawing of conclusions as to the susceptibility of the tube type *6J5WGT* or the 6000 series. A microphonic spectrum taken at sound levels of 145 to 160 db would, no doubt, show marked increase in the noise output of all tubes, especially at the various resonant frequencies of the tube structure. Of particular interest is the fact that some of the worst resonant conditions occur above the frequency range of the Joint Army-Navy specification for vibration testing of electron tubes.

RELAY TESTING

Sensitive relays of the "null-seeking" variety are used in aircraft and guided missiles as the power amplifying output of the guidance or control servos. The effects of shock and vibration on those sensitive components are generally recognized. Not so well recognized, however, is the effect of high-intensity sound. For this reason, sensitive null-seeking relays were included among the components initially investigated in the sound chamber.

A null-seeking sensitive relay consists of two or more control windings, a light armature, and two or more sets of contacts. When the control coils are not energized, the armature is at rest in the center and no electric contact is being made. A small magnetomotive force generated by 1 or 2 mils' control current is sufficient to cause the armature to leave the center rest position and close on one set of contacts.

For purposes of determining the effectiveness of the ARF sound chamber in the testing of this type of relay, a sample from a nationally known manufacturer was tested in the following manner. Provision was made to supply 0 to 5 mils of direct current to the control windings while 1 ampere of direct current was caused to flow through the closed contacts. A thyratron continuity indicator was utilized to detect opening of closed contacts and closing of open contacts, while a d-c oscilloscope was connected across the closed relay contacts to indicate any changes in contact resistance. The relay was mounted within the sound chamber by the same means as would be utilized in normal application.

The relay was subjected to both single-frequency sound fields and white noise, and it was found that 110 db was sufficient to cause disturbance. At certain frequencies, a level of 130 db was sufficient to cause closed contacts to open. The oscillograms of Fig. 10 show variations in contact resistance. If the contact resistance had remained constant, the oscillograms would appear as a straight line. The lower limit of the trace represents the effective zero resistance obtained with maximum contact pressure, while the deviations from the lower limit indicate increasing resistance as contact pressure decreases. At the sharp peaks shown in the record for 1,100 cps excitation and 130-db sound field, the contacts actually opened, as was indicated by the continuity indicator.

Testing at single frequencies with the control coils energized so as to produce

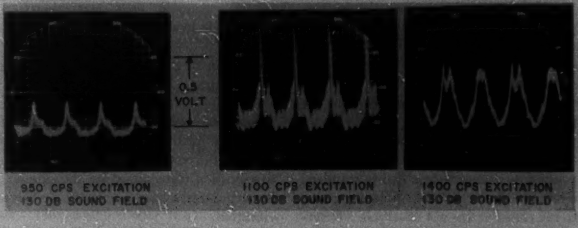


Fig. 10. Relay contact resistance variations in high-intensity sound field

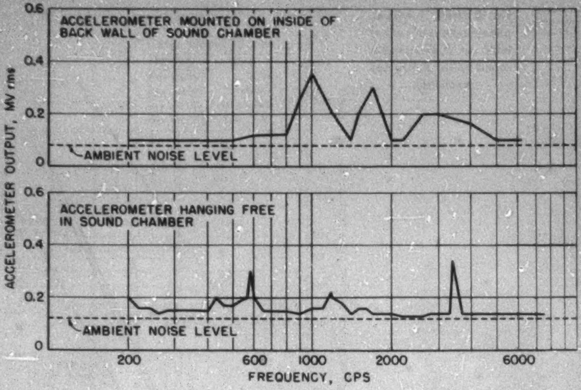


Fig. 11 (above). Accelerometer output in 120-db sound field

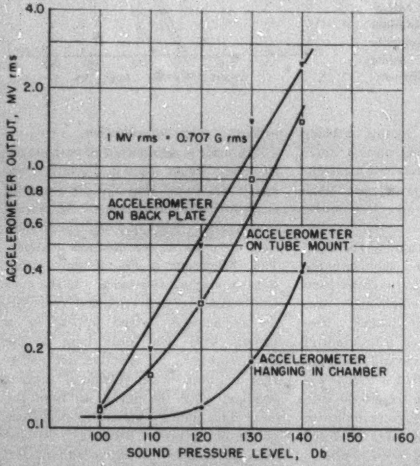


Fig. 12 (left). Accelerometer output at increasing sound pressure levels, white noise excitation

positive closing of one set of contacts indicated a band of resonant frequencies extending from 1,000 to 1,600 cps, and from 2,200 to 2,500 cps. Resonances in the region of 1,000 to 1,200 cps resulted in intermittent opening of the contacts and would thus appear to be the most destructive. It is of interest that the resonances and resultant malfunctions occurred far above the upper limit of frequencies for vibration testing, that is, 55 cps, as specified in military specifications for relays.

Similar results indicating the excitation of relay armatures and changes in contact resistance were noted in standard relay samples tested, but actual opening to the contacts was not observed. Referring to Fig. 1, it may be seen that the sound pressure level of 130 db is not unusual for jet aircraft; consequently, this test should not be considered excessively severe. Sensitive relays, because of their function, represent a more critical problem in their isolation than the standard relays; acoustic tests are, therefore, desirable in evaluating such relays.

Vibration Level Encountered with the Chamber

Since acoustic excitation is a form of vibration testing, the question arises whether the results found are actually due to vibration transmitted by the supporting structures, rather than to the air-borne vibration.

It was felt that the level of mechanical vibration in the chamber was low enough to be considered negligible. To obtain some quantitative data, vibration measurements were made both with single-frequency excitation and with white noise at a sound level of 120 db. Accelerometer output was measured with the accelerometer hanging free in the chamber, mounted on the inside of the back wall, and on the component-supporting bracket. Hanging the accelerometer free in the chamber yields essentially the microphone response of the accelerometer separated from the mechanical acceleration. A Glennite model *A315* accelerometer was used.

Fig. 11 is the output of the accelerometer as a function of frequency when hanging free in the sound chamber and mounted on the back wall. Oscilloscopic examination of the accelerometer signal showed that it was not primarily acoustic pickup of the input disturbance. Rather, it was a combination of contact noise, cathode follower noise and general disturbance. In general, it may be said from Fig. 11 that, at a sound level of 120 db, the vibrational acceleration measured on the chamber back plate does not exceed 0.4 *g* peak over the frequency range of 200 to 6,000 cps and is less than 0.2 *g* peak over the greater portion of that range.

A vibratory acceleration of 0.4 *g* peak is rather low to produce the microphonic output encountered in the acoustic tests. In fact, earlier vibration tests of the *6J5WGT* tube had demonstrated that vibration levels of 10 to 20 *g* were insufficient in some samples to yield microphonic outputs as high as those observed on the 120-db sound field. Consequently, it appears that the results obtained with tubes and other components in the 120-db region were caused mainly by the acoustic excitation and not by secondary vibrational excitation at the same frequency.

The accelerometer output for the accelerometer mounted on the back plate, the component mounting bracket, and hanging free in the chamber, at increasing levels of white noise excitation, is shown in Fig. 12. The acceleration recorded at the mounting bracket at 140 db is 1 *g*.

On the basis of these acceleration measurements, it is felt that the results obtained in acoustic testing are caused mainly, if not solely, by the acoustic excitation. The existence of vibration is recognized, and the causes have been noted. Certain modifications of the chamber intended to increase acoustic efficiency will also tend to decrease the vibration at various sound levels; fore-most among these is the use of much more rigid walls.

Summary and Conclusions

A device for laboratory testing of small electronic components has been designed and built, utilizing readily obtainable electronic equipment and a simply constructed chamber. Instrumentation is relatively simple, and calibration and operation are straightforward. Although the system is still in an early stage of development, sound pressure levels of the order of 150 db have been obtained, utilizing both single-frequency and random noise source inputs.

Preliminary testing of tubes and relays has yielded valid and useful data which point the way toward the need for further investigation of air-borne electronic components in the range above the standard vibration test frequencies.

Potential improvements in the performance of the ARF high-intensity sound chamber may be considered in terms of efficiency of conversion of electric to acoustical energy or of attainment of higher sound pressures. Some obvious approaches toward improvement in either or both of these respects are:

1. Coupling the speaker drivers to the enclosure through a load-matching device, e.g., perhaps a modified horn.

2. Building the chamber with principal dimensions unequal and not related integrally; this will broaden the response curve by introducing additional eigenfrequencies.

3. Building the chamber of heavier walls (as envisaged originally) to reduce radiation losses.

4. Use of variable wall spacing (i.e., one wall movable) to give a resonant situation over a wide range of frequencies.

5. Use of higher energy acoustic sources, such as a siren or an air-modulated speaker (with corresponding design changes in the box).

References

1. EVALUATION OF MECHANICAL DESIGN LEVEL OF ELECTRONIC EQUIPMENT LEADING TO VIBRATION AND SHOCK DESIGN CRITERIA. Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, Contract No. AF33 (616)-223, ARF No. KO44.

2. A FIFTY-HORSEPOWER SIREN, **R. Clark Jones.** *Journal*, Acoustical Society of America, New York, N. Y., vol. 18, 1946, pp. 371–87.

3. A POWERFUL HIGH-FREQUENCY SIREN, **C. H. Allen, I. Rudnick.** *Ibid.*, vol. 19, 1947, p. 857.

4. ACOUSTIC AIR-JET GENERATOR, **J. Hartman.** *Det Hoffenbergske Establissement*, Copenhagen, Denmark, 1939.

5. PERFORMANCE OF EUVES SOUND PROJECTOR. *OSRD Report 1528, Office of Scientific Research and Development*, Washington, D. C., 1954.

6. AN ELECTROMAGNETIC SOUND GENERATOR FOR PRODUCING INTENSE HIGH FREQUENCY SOUND, **H. W. St. Clair.** *Review of Scientific Instruments*, New York, N. Y., vol. 12, 1941, pp. 250–56.

7. VIBRATION AND SOUND (book), **Philip M. Morse.** McGraw-Hill Book Company, Inc., New York, N. Y., 1936, p. 312.

8. DESIGN OF LARGE SHOCK TUBE. Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, Contract No. AF33 (616)-266, ARF No. MO37.

9. MEASURING SHOCK-WAVE PRESSURES, **A. Siegelman, F. Mintz.** *The Frontier*, Armor Research Foundation, Chicago, Ill., March 1954.

10. 6000/CT SERIES ELECTRON TUBES. *Technical Note WCLC 53-11*, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

11. *Military Specification MIL-E-1B*, Department of Defense, Washington, D.C.

◆

Discussion

M. Peila and **G. Freeman** (Boeing Airplane Company, Seattle, Wash.): A 6-inch cube resonant chamber of 1/2-inch aluminum plate driven by a 30-watt drive unit was constructed, and the preliminary data obtained raise the following questions and comments.

When using single-frequency excitation, are the sound levels measured at the microphone position in Fig. 3 indicative of those at the position of the component under test? The sound field distribution within our cube was explored. At most of the discrete test frequencies, approximately 100 to 10,000 cps, the levels varied about 10 db, and, at the chamber resonant frequencies, the variation was found to be higher. The results of these tests indicate that the sound pressure field at the component test position should be known to correlate accurately the results. Random noise excitation was found to provide a more uniform sound field, as pointed out in Mr. Mintz's paper.

Is a vibratory acceleration of 0.4 *g* low for a vacuum tube when considering high frequencies, and is the mounting bracket to be considered typical? Testing has shown that an acceleration of 0.4 *g* applied to the socket of a tube can result in greater acceleration of the tube, and accelerations as low as 1 *g* can produce microphonic outputs far exceeding the *MIL-E-1B* specification (see reference 11 of the paper), particularly at high-frequency resonances even though the *MIL-E-1B* test is met at the test specified frequency. During the testing of vacuum tubes at similar sound pressures we found that the microphonic output was greatly reduced when a large lead block mounting was used, isolated against vibration of the chamber, in place of a simple light bracket. These facts lead us to the conclusion that vibration is significant and dependent upon the mount. Thus, the testing of vacuum tubes in high sound pressures should be done on a package rather than a component basis. This takes into account the mounting peculiar to the package.

F. Mintz and **M. B. Levine**: The questions raised by Mr. Peila and Mr. Freeman are of particular interest because we learned on December 6, 1954, that their work, using a sound chamber similar to ours, was actually done independently of and possibly earlier than our own work. The basic questions at issue regarding the measurement of sound level are: What is a valid measure of the acoustic excitation in the chamber? Is a sound level, measured at a single point, an adequate measure?

There is no doubt that the acoustic field in the sound chamber (with discrete-frequency excitation) is nonuniform. Since the dimensions of the test object are reasonably large relative to the dimensions of the chamber and also relative to the half-wave length of the sound at frequencies above 1,000 cps, measurement of the sound level at any single point cannot describe the acoustic excitation fully. Consequently, our technique of measuring the sound level at a single point near the test object represents only an approximation of the actual acoustic excitation. It served, however, as a satisfactory basis of comparison for a test condition from one test specimen to another.

Strictly speaking, one should know the complete distribution of sound pressures acting on the surface of the test item. Practically, this might be obtained by a systematic series of perhaps 10 or 12 measurements around the test object. Perhaps the most satisfactory approach would be to determine the sound power input to the chamber. This could probably be done by making a fairly small number of measurements at standard locations in the chamber without specific reference to the test object. Possibly the acoustic system, consisting of the sound chamber, driver unit, and amplifier could be calibrated in terms of sound power as a function of electric power input for different frequencies. In subsequent tests, the electric power input could be measured, serving to indicate the acoustic power input.

With respect to the significance of vibration at the mounting bracket and the effect of the mounting bracket itself on the tube vibration, the first point made is that testing (presumably by Mr. Peila and Mr. Freeman) has shown that an acceleration of 0.4 *g* applied to the socket of the tube can produce microphonic outputs exceeding the *MIL-E-1B* specification. Our own tests show that the microphonism obtained with acoustic excitation greatly exceeds that found with vibration excitation of considerably larger amplitude than 0.4 *g*. Details are given in the following.

It is not clear in what way the application of 0.4 *g* to the socket of the tube results in greater acceleration of the tube. This point may constitute merely a problem in semantics or in reference co-ordinates. All of our vibration measurements were made at the tube support bracket. As a matter of convenience, the mounting bracket used for the tubes in the acoustic tests was the same as used for the vibration tests. Consequently, our data comparing vibration and acoustic test results would be completely consistent.

It may be seen from Fig. 11 that the peak acceleration signal obtained was equivalent to about 0.35 *g*; this was obtained at only one frequency (about 1,000 cps) well removed from the principal resonant frequency at which microphonic failure was observed in the tubes. Throughout most of the range, the acceleration level was below

0.2 *g;* and below 600 cycles, most of the signal was actually ambient noise.

Of 24 *6J5WGT* tubes tested in vibration, five samples exceeded the *MIL-E-1B* microphonic level at 1 *g* at the tube resonant frequency; five reached or exceeded the failure level at 2 *g;* one at 3 *g;* two at 4 *g;* one at 5 *g;* and one at 8 *g*. Six tubes tested at 5 *g*, two tubes tested at 10 *g*, and two tubes tested at 20 *g* had microphonic levels less than the *MIL-E-1B* failure level. From the viewpoint of our tests, therefore, it is quite apparent that the vibration level as measured on the tube support did not contribute appreciably to the observed microphonic effects.

It is not surprising that the use of a large lead block isolating against chamber vibration resulted in lower microphonic levels than those obtained with a simple light bracket. Assuming that some internal vibrations of the tube were generated by an acoustic disturbance, this would naturally cause alternating force reactions at the base of the tube. The damping and inertial effect provided by a large lead block at the base would undoubtedly result in lower internal vibrations for a given acoustic excitation.

This does not necessarily justify the testing of vacuum tubes on a complete package basis; it implies primarily that the testing method should reflect in reasonable measure the mounting conditions to be expected in a practical electronic package. The advantages of testing by component rather than by equipment, when it is the component that is of interest, are too obvious to require elaboration here.

Gearing for Diesel-Electric Locomotives

G. T. BEVAN
MEMBER AIEE

TWO means of improving railroad operating efficiency are related to locomotive operation and performance. One is to obtain maximum locomotive utilization, and the other is to increase gross ton-miles per train-hour.

Characteristics inherent in the steam locomotive prevented the exploitation of these two means to the fullest extent. The characteristics of electric drive as applied to Diesel-electric locomotives are more favorable. Consequently, the wide application of this type of motive power has enabled the railroads to advance far toward maximum locomotive utilization, and to increase materially their gross-ton-miles-per-train-hour ratio.

Within the limits of engine horsepower (hp), the hauling capacity of a Diesel-electric locomotive is determined by weight on drivers, traction motor rating, and motor gear ratio. The weight on the drivers of a locomotive permits it to haul a specified tonnage, depending upon the adhesion characteristics between the wheel and the rail. The traction motor rating allows the locomotive to pull this tonnage at a definite speed for a specified length of time, depending upon the profile of the line and operating conditions. The traction motor gear ratio permits the locomotive to be operated in certain classes of service.

The first two factors play a most important part in determining the gross ton-miles that a locomotive with a given hp rating can produce per train-hour. The third largely decides the way in which the locomotive can be applied, since locomotive gearing offers a means of utilizing hp in the desired combination of speed and tractive effort. The larger the traction motor, the more opportunity there is for using higher speed gearing with a minimum sacrifice of tractive effort.

Dual-Purpose Gearing

Traction motors, such as the *GE-752*, are in use today which permit railroads to take advantage of both high tractive effort and high-speed gearing. Where dual-purpose operation is desirable, as in the pooling of locomotives for passenger and freight service, this type of motor is playing a prominent role. For instance, where a railroad desires to have full tractive effort and at the same time have a relatively high-speed locomotive, a 75-mile-per-hour (mph) gearing (65/18 gear ratio) instead of a 65-mph (74/18 gear ratio) can be used without any sacrifice in haulage capacity. This is illustrated by Fig. 1.

The effects of such adverse factors as worn wheels, climatic conditions, type of sand, and rail head shape make it generally unlikely that adhesions over 18.5 per cent can be obtained in day-to-day operation throughout the year.

In the case of a locomotive, the coefficient of adhesion between the wheels and the rail can be expressed as the ratio between the tractive effort that can be maintained up to the slipping point of the driving wheels and the total weight on drivers.

Experience has shown that adhesions higher than 18.5 per cent are not to be relied upon when poor rail conditions such as snow, frost, moisture, leaves,

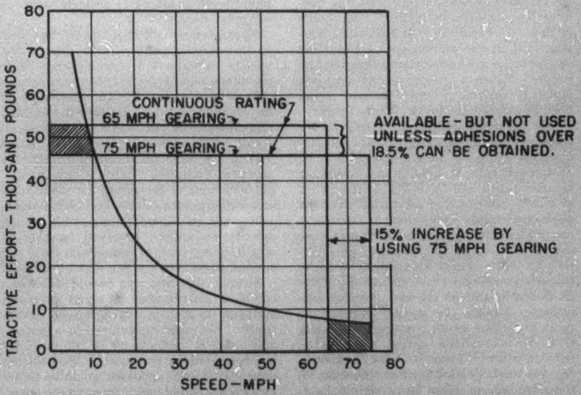


Fig. 1. Speed-tractive effort curve for Diesel-electric locomotive. 65- and 75-mph gearing, 1,600 hp, four axles, 125-ton unit, 40-inch wheels, four GE-752 motors

---

Paper 54-382, recommended by the AIEE Land Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing August 16, 1954.

G. T. Bevan is with the International General Electric Company, New York, N. Y.

Table I. Comparison of Tractive Efforts, 75- Versus 65-Mph Gearing

1,600-Hp 4-Motor 4-Axle Diesel-Electric Locomotive, 62,500 Pounds per Axle, 40-Inch Wheels

	75-Mph Gearing	65-Mph Gearing
Gear ratio	65/18	74/18
Tractive effort at continuous rating, pounds	46,500	53,000
Adhesion at continuous rating, per cent	18.6	21.2
Tractive effort at 18.5-per-cent adhesion, pounds	46,250	46,250
Motor current limitations at 18.5-per-cent adhesion	continuous	continuous

Table II. Comparison of Tractive Efforts, 75- Versus 65-Mph Gearing

2,250-Hp 6-Motor 6-Axle Diesel-Electric Locomotive, 62,500 Pounds per Axle, 40-Inch Wheels

	75-Mph Gearing	65-Mph Gearing
Gear ratio	65/18	74/18
Tractive effort at continuous rating, pounds	69,800	79,500
Adhesion at continuous rating, per cent	18	21
Tractive effort at 18.5-per-cent adhesion, pounds	69,400	69,400
Motor current limitations at 18.5-per-cent adhesion	continuous	continuous

etc., are encountered. Wheel tread wear, low rail joints, severe curvature, rail oilers, and special trackwork also may cause slipping. Whatever the cause, slipping results in reduced tonnage or loss of time, or both. Sand is used to counteract poor rail conditions. When of suitable quality and correctly applied, it is usually effective.

Under the average operating conditions just described, the advantage of 75-mph gearing, such as that offered on the GE-752 motor, is made clear by comparing the data in the two columns of Table I. These are based on Diesel-electric locomotives of the 1600-hp size weighing 125 tons, all on drivers.

From the figures in Table I it is evident that the 75-mph gearing permits greater utilization of the capabilities of the electric traction equipment. In other words, full use is made of the available continuous tractive effort without exceeding normal operating adhesion values.

In average freight service either the 65-mph or the 75-mph locomotive could pull the same tonnage over a given profile and, therefore, could accumulate the same gross ton-miles per train-hour for a particular run. But the locomotive geared for 75 mph could also handle passenger trains on a turn-around basis, thus raising the locomotive utilization factor. On the other hand, the locomotive geared for 65 mph is limited to a somewhat lower speed service. This limitation has a tendency to increase the number of locomotive units required to handle an assignment involving both freight and passenger trains. As a result, locomotive investment may be increased and unit mileage decreased, tending to produce higher operating costs per locomotive-mile.

When considering the case of a 2,250-hp 6-motor locomotive with 65 and 75-mph gearing, Table II outlines a similar comparison. As indicated in this table, the unit with 75-mph gearing utilizes all of the available continuous tractive effort at 18.5-per-cent adhesion. If track conditions permit adhesions on the order of 20 per cent, the use of short-time traction motor ratings could be considered. Here again, as in the case of the 1,600-hp unit, the higher speed gearing permits a wider application of the locomotive and gives a correspondingly increased potential utilization.

Application Example

In selecting the correct gearing for a given locomotive application, the fundamental requisite is to determine the general type of usage involved. For example, a class-I railroad desires to handle both freight and passenger service with one type of locomotive unit. The basic data considered for this example are:

Length of division, 110 miles.

Length of level tangent track, 100 miles.

Length of 1-per-cent grade, 10 miles.

Freight train speed desired on level track, 40 to 50 mph.

Freight train speed up grade, not to fall below 10 mph.

Desired running time for freight trains, not over $3^{1}/_{2}$ hours.

Maximum permissible passenger train speed, 75 mph.

Locomotive unit specified by railroad, 1,600 hp.

A comparison of the two gear ratios that might be recommended for this particular service is given in Table III.

In considering passenger service, it is assumed that the locomotive operating in freight service in one direction returns with a passenger train consisting of ten 70-ton cars. With this type of assignment, the locomotive geared for 75 mph would permit faster train schedules than the one geared for 65 mph. The difference in running time over the division would amount to 6 minutes.

The ratio of gross ton-miles per train-hour is a handy yardstick for measuring operating efficiency. Today, every railroad is attempting to lower operating costs by increasing this ratio. For example, over the past 10 years, it has been rising at the rate of approximately 4.5 per cent per year. The present trend appears to be demanding more

Table III. Locomotive Application

1,600-Hp 4-Motor 4-Axle Diesel-Electric Locomotive, 62,500 Pounds per Axle, 40-Inch Wheels

	Column 1	Column 2
Locomotive hp	1,600	1,600
Maximum permissible speed, mph	75	65
Gear ratio	65/18	74/18
Tractive effort at continuous rating, pounds	46,500	53,000
Adhesion at continuous rating, per cent	18.6	21.2
Tractive effort at 18.5-per-cent adhesion, pounds	46,250	46,250
Freight Service		
Trailing tons (50-ton cars, limited by 1-per-cent grade)	1,750	1,750
Speed on level track	39.5 mph	39.5 mph
Speed on 1-per-cent grade	10.0 mph	10.0 mph
Running time, 100 miles level track†	2 hours 33 minutes	2 hours 33 minutes
Running time, 10 miles 1-per-cent grade, hours†	1	1
Ton-miles per trip	192,500	192,500
Ton-miles per train-hour	54,250	54,250
Passenger Service		
Trailing tons (ten 70-ton cars)	700	700
Speed on level track, mph	68.5	65*
Speed on 1-per-cent grade, mph	25.5	25.5

* Limited by locomotive gearing.

† Acceleration and braking times neglected.

dual-purpose locomotive operation. The better motive power utilization achieved in this way means that not only are less locomotive units required but also less facilities are needed for inspection and maintenance.

It must be remembered, however, that Diesel-engine hp alone will not produce the desired results. There must be a corresponding increase in traction motor capacity and a proper choice of gear ratio; otherwise the locomotive will not be able to produce the additional ton-miles per train-hour indicated by its higher hp rating.

Advantages of Dual-Purpose Gearing

In view of the foregoing considerations, it appears that a higher gear ratio will tend to produce a more versatile locomotive. Such a unit could be expected to have greater potential earning capacity and lower operating cost per mile. There are additional reasons why the 75-mph gearing is a potential means of reducing operating costs as compared to the 65-mph gearing.

1. Armature rotating stresses are reduced.
2. Commutation is improved.
3. Commutator maintenance and brush life are improved.
4. Incidence of flashover is reduced.
5. In the case of the *GE*-752 motor, the use of higher speed gearing reduces the armature speed corresponding to a given locomotive speed by approximately 12 per cent.
6. In the case of the 1,600-hp locomotive under consideration, a wheel slip on any one axle will tend to stabilize at the same speed with either 65-mph or 75-mph gearing. The percentage of armature overspeed, however, will be considerable lower on the 75-mph geared locomotive.

Conclusion

In the earlier days of Diesel-electric locomotive design and application, traction motor gear ratios which would best suit the locomotive to a specific type of service were selected. As engine hp and traction motor capacity have increased, it has become more and more desirable to select gear ratios which will permit greater locomotive utilization by enabling the same locomotives to be employed in either freight or passenger service and, in many instances, in switching service also. The benefits resulting from this change are threefold. the number of locomotives for a given assignment is reduced; the operating costs, which are directly affected by locomotive mileage, are decreased, and the requirements for maintenance facilities are also reduced. All of these results contribute to more efficient railroad operation.

◆

Discussion

H. F. Brown (Gibbs & Hill, New York, N. Y.): The author has presented a very convincing paper on the advantages of a higher gear ratio for Diesel-electric locomotives. Much stress has been laid during the recent past on the prevalence of wheel slip on Diesel-electric locomotives. Tractive-effort ratings have been based generally on 25-per-cent adhesion, and the lower gear ratios have enabled adhesions of these values to be closely indicated at continuous ratings. Another paper[1] also indicates that attempts to realize such high adhesion values are a large factor in causing wheel slip.

As Mr. Bevan points out, however, adhesions of 25 per cent can rarely be realized in everyday operation. It would appear, therefore, that one of the greatest advantages to be obtained by the use of a higher gear ratio would be in the reduction of wheel slip and in some of the flashovers which often follow such slipping.

The interesting moving pictures of generator flashovers presented at this meeting show "ring fire" caused by flashing being blown away by the application of a stream of compressed air. This has inspired the thought that possibly generator flashovers might be reduced by the judicious application of a small stream of compressed air directed at the commutator so as to neutralize the air currents following the direction of rotation. This might tend to keep any ionized gases from bridging across between brush holders.

It has been common practice for years to use power to blow air through the traction motors for their cooling. A little of the compressed air always available on a locomotive might be used for such protection.

Reference

1. Diesel-Electric Locomotive Wheel Slipping; Causes, Effects and Methods of Control, R. I. Fort. *AIEE Transactions*, vol. 73, pt. II, 1954 (Jan. 1955 section), pp. 478–80.

Development of a Large Ratio-Release Circuit-Breaker Mechanism

J. A. SCARCELLI
ASSOCIATE MEMBER AIEE

RUDOLF STEINER
ASSOCIATE MEMBER AIEE

DESPITE commendable efforts and admirable successes toward the development of aircraft circuit breakers, a wholly acceptable solution apparently has not yet been found; therefore, much can be gained from the observation of existing imperfections. In addition to new and original creations, considerable contributions toward improved circuit-breaker constructions can be had through the process of elimination applied to present shortcomings. It will aid in the examination of circuit-breaker capabilities to recognize the circumstances which prompted the earlier adoption of specific features which are now undesirable. Various reasons which may account for these conditions are: lack of earlier requirements, lax co-ordination among protection devices and circuit-equipment specifications, or excessive speed applied to the development of certain circuit breakers because of national emergencies.

In so far as the circuit breakers designed under conditions considered normal are concerned, a number of characteristics can bear both ample modification and improvement. Disregarding one specific variety that does not utilize a mechanism, it became apparent that either vital properties were not incorporated or certain circuit-breaker elements were required to perform more duties than could reasonably be expected from them.

To design and develop systematically a circuit breaker, it seems necessary to segregate the various problems and to attack each element individually. Obviously, the mechanism, which is the heart of a circuit breaker, deserves first consideration. The purpose of this paper is to present a methodically developed circuit-breaker mechanism allowing for

---

Paper 54-423, recommended by the AIEE Air Transportation Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 9, 1954; made available for printing August 27, 1954.

J. A. Scarcelli and Rudolf Steiner are with the U. S. Naval Air Development Center, Johnsville, Pa.

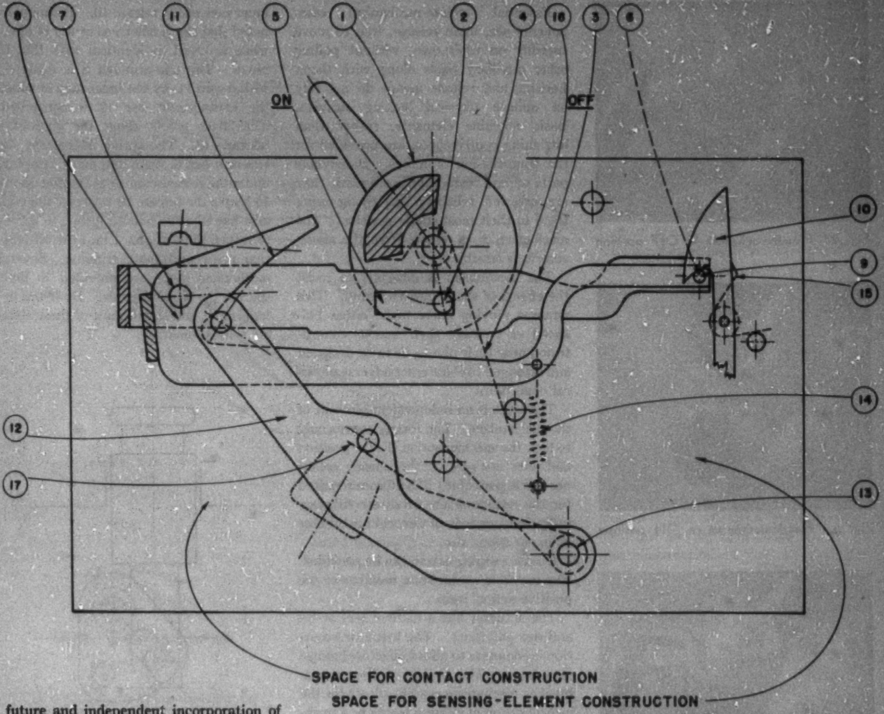

8
7
11
5
1
2
4
16
3
6
ON
OFF
10
9
15
14
12
17
13
SPACE FOR CONTACT CONSTRUCTION
SPACE FOR SENSING-ELEMENT CONSTRUCTION

Fig. 1. **Large ratio-release mechanism**

——— Visible contours
— — — Invisible contours
- - - - - Mechanical springs
—·—·— Center lines
—··—··— Suggested contours

future and independent incorporation of contact arrangements as well as for arc suppression and sensing means. In addition the mechanism was designed to allow for any desirable actuator variety without modification of construction. Finally, the design was prepared with the intent to permit ready conversion from single- to multi-pole units with the least amounts of additional parts.

Existing Conditions

Field reports of unsatisfactory circuit-breaker performance indicate a variety of offensive conditions found with existing circuit breakers. Among these, the following appear to be the most prominent:

1. The actuator is an integral member of the mechanism and is not readily replaceable with an actuator of another more desirable or suitable variety.

2. The sensing unit, particularly that of the thermal type, is a part of the mechanism or of the release.

3. The movable contacts drag a part of or the whole mechanism along, thus retarding considerably the vital contact opening speed.

4. Many circuit-breaker mechanisms engage the movable contact arm at a point near its pivot, thus establishing an undesirable cantilever condition, i.e., unreliable locking qualities in the closed (ON) position.

5. A number of circuit-breaker varieties do not provide "positive make and break"; some circuit-breaker types offer only one or the other of the two features.

6. The majority of the known circuit-breaker constructions have unbalanced movable parts which assume an unreliable attitude. This is true particularly under conditions of shock or vibration. Others have artificially counterbalanced parts, thus adding dead weight to the assembly.

7. The conversion, if at all possible, of several known single-pole circuit breakers to multi-pole units would require major redesign.

8. The release portion of most known circuit breakers is usually an intricate and delicate detail requiring unreasonably small tolerances of the parts involved. In view of this circumstance and the undesirable cantilever arrangement mentioned in item 4, these component parts are subjected to excessive wear and tear resulting in erratic performance or untimely malfunctioning.

Definition of Specific Terms

The following are definitions of specific terms used throughout this paper:

"Large ratio-release circuit-breaker mechanism" concerns the release arrangement which is large in comparison to the movable-contact member controlled by it. This effect was primarily attributable to the utilization of a second-class lever.[1]

"Positive action" characterizes the operational property of the circuit-breaker mechanism amounting to a substantially continuous, physically inflexible engagement among all mechanical parts involved.

Accomplished Objectives

Upon analysis of objectionable circuit-breaker characteristics and consideration

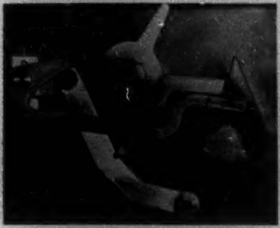


Fig. 2. Pseudomechanism in OFF position

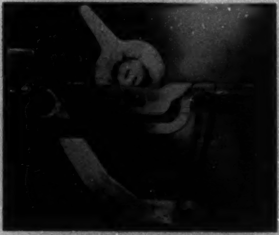


Fig. 3. Pseudomechanism in ON position

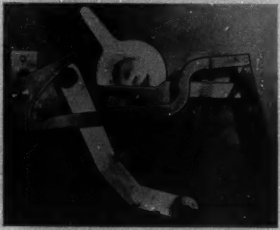


Fig. 4. Pseudomechanism in TRIPPED position

of those properties which would aid in the satisfactory performance of the circuit-breaker mechanism, the following design objectives become apparent in regard to the general assembly, the mechansim proper, the resistance to shock and vibration, the latch, the contact action, and the actuator.

The assembly is straightforward without relying on unreasonable tolerances. Further, it does not require specialized facilities on the part of the assembly personnel.

The mechanism is convenient for inspection in each operating position. The movements are fundamental, i.e., they utilize straight-line and circular motions. Any complex motions are avoided. Correlated movements are coacting.

Essential release-mechanism components are, upon release, free to move speedily on their own, without pulling other auxiliary parts along with them. Locking and release means do not rely on minute physical locking surfaces. Basic machine elements, rather than intriguing contrivances, are provided for the transmission of forces. If movable parts require temporary positioning, they are arrested reliably, i.e., by securing them in their most stable position. The mechanism is as remote from the arcing zone as is feasible.

The mechanism is designed to resist the effects of shock and vibration. This requires moving parts to be either balanced or to have their masses concentrated near their pivots, or to be designed and arranged to counteract adverse physical conditions.

The latch is an independent element of the mechanism. The former serves only to lock the mechanism on the ON position and does not rely on the sensing means nor is it a part of it. The forces required for the release of the latch are substantially independent of the circuit-breaker rating or frame size.

Contact wiping action can be provided. The movable contact is a member of the positive-action train.

The actuator has a distinct feel at ON and OFF switching. The actuator excursion is adequate to afford effective linkage movement. The design of the circuit-breaker mechanism readily allows for the incorporation of either a toggle or a push-pull type of actuator, as required by the particular application. The actuator is a member of the positive-action train.

The design of the circuit-breaker mechanism permits convenient adaptation from the basic single-pole unit to multi-pole units.

Operation of the Mechanism

DESCRIPTION

The parts of the circuit-breaker mechanism are identified in Fig. 1. A pseudomechanism, shown in Figs. 2, 3, and 4, will aid in the observation of the operation. The members in Fig. 1 show the construction in front elevation and in cross section. The actuator 1 is pivoted about pin 2. This pin is located within the circuit-breaker case. The actuator engages the actuator bar 3 through a pin 4 operating within the slot 5 of bar 3. The latter is pivoted about the point 6 which is positioned within the circuit-breaker case. About pin 7 secured to bar 3, the trip-arm bar 8 is arranged free to rotate unless its pin 9 is arrested beneath the inner race of the release 10. It should be noted that the geometrical center of pin 9, while arrested, is identical with that of pin 6. The trip-arm bar 8 is equipped with a cam 11, by the inner race of which the contact-arm bar 12 is controlled. The latter pivots about the stationary fulcrum 13. The spring 14 returns the trip-arm bar 8, upon automatic tripping, under the release cam 10. Another spring 15 biases the release 10 to arrest the trip-arm bar 8. The handle spring 16 serves to return the actuator 1 into the off-position upon automatic tripping. Several mechanical stops are provided to limit specific parts movements. Both the arrangement and significance of these stops are self-explanatory.

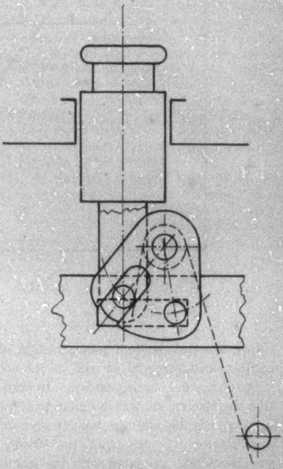


Fig. 5. Push-pull type of actuator

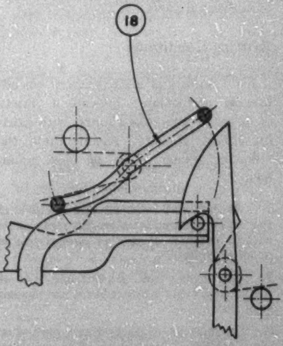


Fig. 6. Multi-pole trip arrangement

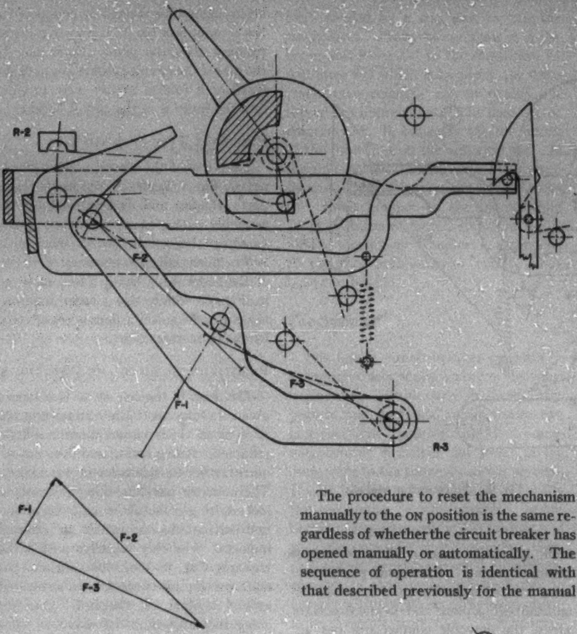


Fig. 7. Force vector analysis

Fig. 1 shows the circuit-breaker mechanism in the ON position. To return it manually into the OFF position, the actuator 1 is turned clockwise, thereby raising actuator bar 3 and, together with it, the trip-arm bar 8. This clears the way for the free travel of the contact-arm bar 12 by virtue of its return spring 17. Attention is called to the feature that, at this operation, the trip-arm bar 8 coacts with the actuator bar 3. Both these bars rotate about the identical centers of pin 9 and 6 respectively.

On automatic tripping, a sensing unit engages the release 10 to cause its clockwise movement and the escape of the trip-arm bar 8 from beneath the circular, inner race. During this operation, the centers of pin 9 and pin 6 do not remain aligned. This establishes clearance for the contact-arm bar 12 to open the contacts (not shown). Subsequently, the actuator bar 3 is no longer under pressure and the handle spring 16 will return the actuator 1 into the OFF position, thereby raising the actuator bar 3. This is followed by the relatching of trip-arm bar 8 underneath the release 10 to restore the mechanism for operation to the ON position.

The procedure to reset the mechanism manually to the ON position is the same regardless of whether the circuit breaker has opened manually or automatically. The sequence of operation is identical with that described previously for the manual opening; however, the parts movements occur in a direction reversed with respect to the action discussed for the manual opening.

The trip-free feature is accomplished by the dual linkage, consisting of the actuator bar 3 and the trip-arm bar 8. Provided the actuator 1 is arrested intentionally in the ON position and the sensing unit engages the release 10 in a clockwise direction, the trip-arm bar 8 can escape, independently, from the still locked actuator bar 3. As in the case of automatic tripping, pin 9 completes its excursion because of the movement of the trip-arm bar 8 while pin 6 remains stationary. The contact-arm bar 12 is free to travel and to open the electric contacts. Upon release of the actuator 1, the parts movements, restoring the manual OFF position of the circuit breaker, occur in accordance with the aforementioned operations.

Coacting Movements

Evidently, the most trouble-free conditions are possible if correlated parts movements are coacting. The following component groups are typical examples of this principle and its application throughout the construction of the circuit-breaker mechanism:

1. The actuator bar has the identical geometrical movement center as the trip-arm bar.
2. The inner cam of the release has a true

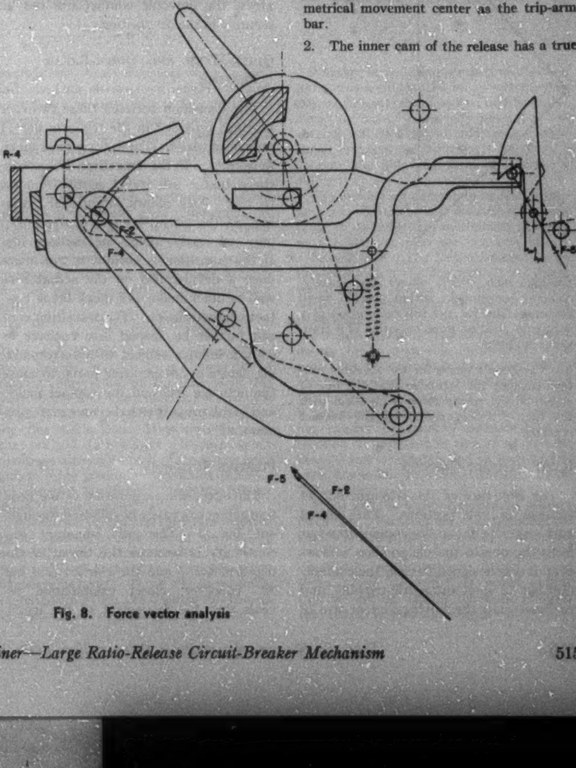
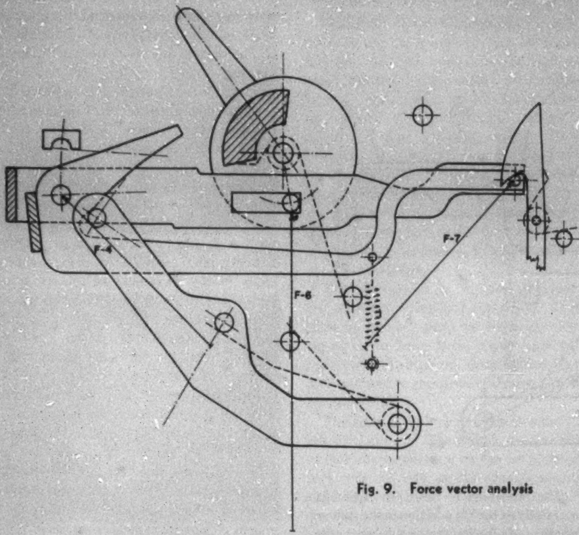


Fig. 8. Force vector analysis



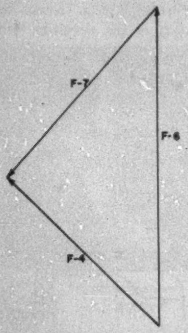


Fig. 9. Force vector analysis

circular race which allows for sensing-element movements within specific small responses and for the full restoration of a secure ON position upon cessation of a transient condition.

3. The return spring for the movable contact engages the latter so that the line of action of the spring force is identical with the line of force exerted by the contact pressure.

RAPID CONTACT OPENING

The cam race of the trip-arm bar 11 consists of two portions. The curved part serves to move the contact-arm bar from the OFF to the ON position. However, it was designed to clear the contact-arm bar 12 upon automatic tripping, and to disassociate the moving contact-arm 12 from its mechanism to allow for rapid contact opening. The second part of the race, located adjacent to the throat, is substantially a plane and its function is to arrest the movable contact-arm bar 12 securely in the ON position.

QUICK-MAKE AND QUICK-BREAK FEATURE

The transition between these two race portions provided in the trip-arm bar 11 may be of various shapes to produce specific effects. If the transition is smooth, the actuator will travel gently from ON to OFF. This allows intentional "teasing" of the circuit-breaker contacts and may adversely affect the contact life. If the transition is abrupt, it will introduce a definite feel on the actuator as well as quick-make and quick-break contact characteristics. The transition zone can further be shaped into concave or convex humps, subject to a corresponding design of co-operating parts, to cause the actuator feel and the contact make-and-break properties to become more pronounced.

POSITIVE ACTION

The effected incorporation of the positive action can easily be followed throughout Fig. 1. The only apparent discontinuity is between the throat of the trip-arm bar 8 and the contact-arm bar 12. However, closer examination reveals that the movement of the mechanism cannot be started or completed if the contact-arm bar 12 is prevented from rotating about its pivot. Such may be the case if the circuit-breaker contacts are fused or if foreign matter were to enter into the interior of the circuit breaker.

PUSH-PULL TYPE OF ACTUATOR

The view of the circuit-breaker mechanism, Fig. 1 illustrates the fundamental parts, shapes, and designs used for the assembly. The few, simple steps required to equip the circuit-breaker mechanism with a push-pull type of actuator, in lieu of the toggle-type handle, are portrayed in Fig. 5, which also clearly indicates that the mechanism proper is not affected by such a modification.

RESISTIVITY TO SHOCK AND VIBRATION

The mass of the release 10 is balanced about its fulcrum. The contact-arm bar 12 is at all times biased through a comparatively strong spring and does not require further consideration in this respect. The actuator bar 3 is either released or locked by the actuator and, therefore, not particularly susceptible to external impacts. The only sensitive part is the trip-arm bar 8. For this reason, the main portion of its mass, is concentrated around its pivot, i.e., the pin 7. The foregoing indicates that the movable parts are not adversely affected by exposure of the mechanism to shock and vibration. It will be noted that the only part arrested in only one direction is the trip-arm bar 8 which rotates clockwise, if at all, and thus provides a more secure locking for the contact-arm bar 12.

MULTI-POLE CONSTRUCTION

The conversion from the single-pole unit to multi-pole mechanisms is illustrated in Fig. 6. This figure shows the elementary arrangement employing one additional part, aided by a return spring which is common for all poles used in the aggregated assembly. The cradle 18, pivoted within the circuit-breaker case, is engaged at one end by the trip-arm bar which escapes from its release upon completed sensing of overload. The other end of the cradle actuates the releases of all other poles of the multi-pole assembly. This causes practically simultaneous tripping of all circuit-breaker poles.

MANUFACTURING TOLERANCES

Throughout the description of the circuit-breaker mechanism, parts, and figures, it becomes apparent that no abnormally close manufacturing tolerances are required for the adequate functioning of the circuit breaker. This is particularly true for the effective multi-pole trip arrangement which can normally be had only by using accuracies peculiar to the field of instrument mechanisms.

Force Vector Analysis

Figs. 7, 8, and 9 indicate the results of a force vector analysis of the circuit-breaker mechanism. The vector study starts with a unit force *F-1* which is considered the total force acting on the movable-contact arm. This force consists of the contact pressure, provided by means of the particular manufacturer's choice, and the force produced by the main return spring. The assumption is made, for this analysis, that both these forces are acting along the same line and such a condition can, without difficulty, be provided in the design. Fig. 7 shows the analysis of the forces acting on the movable contact arm. The force *F-2*, acting from the contact-arm bar on the trip-arm bar, is resolved into two components along the trip-arm bar; see Fig. 8. It is the force *F-5* acting from the trip-arm bar on the release that the trip unit must supply to operate or trip the circuit-breaker mechanism, except that the force *F-5*, shown in Fig. 8, has not been considered with the friction coefficient to indicate the smaller resultant force needed by a trip unit to trip the mechanism. From the comparison of the given unit force *F-1* to the magnitude of trip force *F-5* which is required, the term "large ratio release," can be readily observed and appreciated. The final vector diagram, Fig. 9, indicates the resolution of the force *F-4* on the actuator bar. It should be noted that the force *F-6* exerted by the actuator bar on the handle locks the circuit breaker on the ON position. The resolution of forces, shown in Fig. 9 for the toggle-type actuator, is applicable to the push-pull actuator variety illustrated in Fig. 5.

Conclusions

The entire circuit-breaker industry is a comparatively young branch of electrical engineering. Quite early in its life, this branch was split into substantially two factions; one favoring the electromagnetic-sensing principle, the other giving perference to the thermal-sensing precept. It soon became evident that the naturally fast-responding electromagnetic circuit breakers required artificial retarding and that the naturally slow-responding thermal circuit breakers needed intentional speedup of their respective responses to comply with the still controversial objectives of circuit protection. It is believed that a great amount of engineering effort was invested in the modifications of the sensing elements at the expense of the more vital circuit-breaker component, the mechanism. It seems fair to state that in many instances the mechanism appears to have been built to suit a specific sensing means, in opposition to what would normally be expected.

To develop and design a circuit-breaker mechanism methodically, it was considered imperative to disassociate it from all other missions without disregarding the requirements to accomplish them. Thus, the presented mechanism evolved without prejudice toward any circuit-breaker components served by or controlling it. The connoted solution appears to live up to the expectations stipulated, particularly for military aircraft circuit breakers. This was accomplished through systematic design of each part, avoiding any disturbing ballast. In this manner, the circuit-breaker construction, intended to be specialized, is readily suitable for less strategic uses such as for industrial, commercial, and residential applications.

Reference

1. MECHANICAL ENGINEER'S HANDBOOK, R. T. Kent. John Wiley & Sons, Inc., New York, N. Y., 1947, "Design, Shop Practice," para. 8-03.

Discussion

D. W. Exner (Boeing Airplane Company, Seattle, Wash.): The authors have made an interesting study of a latch mechanism divorced from the problems of contact structure and overload sensing mechanism design. Unfortunately, these two functional mechanisms have very real design requirements of their own. For instance, it has become increasingly evident that aircraft fault-clearing requirements are difficult to meet without double-break contacts, well-separated from each other. Sensing mechanisms of different types have their own advantages, disadvantages, and problems.

Now, all of these different basic mechanisms have to be crammed into a small case having standardized outline dimensions. To provide more space could require building larger airplane cockpits, for there may be 50 to 150 circuit breakers which must be within reach of the flight crew. If it should be found necessary to increase circuit-breaker dimensions to achieve better performance, an increase in the depth perpendicular to the mounting surface would be the least objectionable. Increases in either of the other two over-all dimensions would be highly objectionable to the aircraft designer.

In defense of the circuit-breaker manufacturers, I know that the problem of space has forced certain compromises in design. It is evidence of their ingenuity that we now have several circuit-breaker designs which accomplish the fault-interrupting requirements of present-day systems with a fair degree of success. It is encouraging to note that development of more capable breakers is continuing, and the excellent design objectives given by the authors should be encouraging to this program.

The requirement for positively separable contacts is no doubt good philosophy, although we seldom have welded contacts on the recent designs except under quite unusual fault conditions. In such cases, it is questionable whether a man could get the breaker open manually before something fused.

Thomas J. Martin (Boeing Airplane Company, Seattle, Wash.): First I would like to express appreciation for the interest in this very timely and important problem shown by Mr. Scarcelli and Mr. Steiner in developing this mechanism. It is certainly true that we still do not have adequate circuit breakers, although much progress has been and is being made. Independent thinking by men like Mr. Scarcelli and Mr. Steiner should be encouraged.

It is further agreed that the forces required to be exerted by the release mechanism, especially in a thermal breaker, should be very small, while the contact pressure must, of course, be large. This is a basic weakness of designs in which the tripping mechanism is required to carry the contact pressure.

Whether or not the suggested mechanism actually is a substantial improvement over existing designs can, of course, best be determined by tests. I sincerely hope that the authors will be able to build some sample circuit breakers and try it out. However, I do note the following possible sources of difficulty.

The mechanism appears to be fairly complicated. I wonder how difficult it is going to be to assemble this mechanism plus the sensing element plus the contacts in the *MS25017* case. The forces required to be exerted by part 11 upon part 12 in closing would appear to be enormous if adequate contact pressure and release spring forces are to be used.

I wonder about the ease of incorporating a double break with this mechanism. Our experience with circuit breakers and circuit-breaker tests to date has consistently served to substantiate our belief that a double break is necessary to achieve adequate interrupting capacity. It must be remembered that recovery voltages can be as high as 50 volts in a 30-volt d-c circuit and as high as 350 volts line-to-line in a 208/120-volt 3-phase 400-cycle system.

What are the forces required to operate latch 10? These forces can probably be readily obtained from a solenoid, but how does the author plan to obtain these forces, combined with adequate travel, from thermal-type elements? It would appear that we are up against the same problem that occurs in existing designs.

J. A. Scarcelli and **Rudolf Steiner:** Mr. Exner's comments on the significance of double-break contacts were highly appreciated. In fact, their future utilization has been contemplated for at least one additional reason, i.e., to preclude the use of the undesirable stranded wire "pigtail" leads which may become responsible for many operational troubles. However, regardless of the future contact-type selection, no attempt has been made to indicate a definite statement on contacts within the expressly outlined development of a mechanism.

It is admitted that advanced circuit-breaker designs may rarely encounter welded contacts. To go a step further, some circuit breakers may never even be called upon to interrupt a circuit on overload. No matter what the future missions of circuit breakers may be, they should be equipped, like all safety devices, with all available means to operate dependably under any unexpected fault condition or any combination thereof. While it may be impossible to open a circuit breaker having solidly fused contacts, this fact may be sufficient to indicate readily, through the built-in positive action, the reason for and the location of the trouble.

In regard to Mr. Martin's apt remarks on the verification of the force relations through tests on sample circuit breakers, it is stated that such a task has been scheduled at the time this response was prepared. The authors may be contacted for correspondence on the test results as they become available.

It was surprising to note that the mechanism was considered to be complicated, since the use of only three structural members seems to approach a minimum of complexity. The problem of crowding the mechanism together with the contacts and sensing means into the small, awkwardly shaped envelope required by the Standard Drawing *MS25017* is a great one and considerable efforts are being expended to accomplish this in the design.

The concern expressed over the magnitudes of forces exerted upon part 12 and, subsequently, upon the release, part 10, will be found to be unwarranted upon review of the vectorial force analysis and the references in the text. The magnitude of the small force *F-5* acting upon the release, part 10, is indicated in the vector diagram shown below Fig. 8. The force required to operate part 10 equals that force *F-5* multiplied by the coefficient of friction between those two metal parts amounting, conservatively, to a factor of 0.25. This resulted in a force vector too small to be shown in the drawings. The unit force *F-1* was assumed to be the vectorial sum of he forces available from both the contact-return and contact-pressure springs. The suggestion for consideration of double-break contacts was both recognized and elaborated on in the foregoing response to Mr. Exner's discussion.

Evaluating the Effect of Nonlinearity in a 2-Phase Servomotor

W. A. STEIN
MEMBER AIEE

G. J. THALER
ASSOCIATE MEMBER AIEE

Synopsis: Equations are derived describing the variation of the gain constant and the time constant of a 2-phase servomotor as affected by the magnitude of the control field voltage. These parameters are evaluated from several simple measurements in the laboratory and are then verified by measuring the time constant experimentally. The use of these nonlinear characteristics in the analysis and design of feedback control systems is discussed.

IT IS common practice to represent a 2-phase servomotor and load with a linear transfer function equation of the type

$$KG=\frac{K}{S(S\tau+1)} \tag{1}$$

An equation of this type is a reasonable approximation of the motor characteristics despite the fact that the 2-phase motor is inherently a nonlinear device.[1] In this paper it will be shown that the nonlinearity of the motor causes a simultaneous variation in the parameters K and τ of equation 1; methods will be described for measuring these variations approximately and a procedure is suggested for applying this information to the analysis and design of servomechanisms.

Fig. 1 shows measured speed-versus-torque curves[2] for a Kearfott type *R-111-2AB*, 2-phase servomotor. It is seen that the lower portions of these curves are straight lines and normal practice extrapolates these portions, as shown by the dotted lines. The result is a family of straight-line speed-torque curves which are used as a linear approximation for the motor characteristics. It is also normally assumed that the straight lines are displaced by equal increments. Since the manufacturer normally gives the speed-versus-torque curve for rated voltages, the straight parallel line approximation can be constructed from the manufacturer's data. It is readily shown that the linear transfer function of equation 1 can be derived from such curves, and the parameters K and τ evaluate as

$$K=\left|\frac{\partial\omega}{\partial V_c}\right| \tag{2}$$

$$\tau=\frac{J}{\left|\dfrac{\partial T}{\partial\omega}\right|}\times(2.590\times10^{-3}) \tag{3}$$

where

ω = speed, radians per second
V_c = control field voltage
J = polar moment of inertia, ounce-inches^2
T = torque, ounce-inches

Equations 2 and 3 also assume that viscous friction is negligible.

Determination of Gain and Time Constant in the Transfer Function

Since Fig. 1 shows that the motor speed-versus-torque curves are not straight parallel lines, it is obvious that K and τ, as defined by equations 2 and 3, cannot be constants. It is normally true, however, that a 2-phase servomotor seldom operates at high speeds and, therefore, the only portion of the curves to be considered is the lower speed portion below 50-per-cent rated speed in which the curves are essentially straight lines, though not parallel.

For this lower speed portion the equation of the family of curves is given by the following

$$T=AV_c-\left|\frac{\partial T}{\partial\omega}\right|\omega \tag{4}$$

where A is a constant and is determined from stalled torque data

$$A=\frac{\text{Stalled torque at rated voltage}}{\text{Rated voltage}} \tag{5}$$

The $\left|\dfrac{\partial T}{\partial\omega}\right|$ term is the reciprocal of the slope of the speed-versus-torque curve and is commonly referred to as the motor damping coefficient.

From equation 4, assuming that viscous friction is present and applying normal mathematical methods, the transfer function of motor and load is

---

Paper 54-522, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted May 17, 1954; made available for printing September 13, 1954.

W. A. Stein and G. J. Thaler are with the U. S. Naval Postgraduate School, Monterey, Calif.

$$KG=\frac{\dfrac{A}{\left|\dfrac{\partial T}{\partial\omega}\right|+F}}{S\left(S\dfrac{J(2.590\times10^{-3})}{\left|\dfrac{\partial T}{\partial\omega}\right|+F}+1\right)} \qquad (6)$$

where F=coefficient of viscous friction. From equation 6

$$K=\frac{A}{\left|\dfrac{\partial T}{\partial\omega}\right|+F} \qquad (7)$$

$$\tau=\frac{J}{\left|\dfrac{\partial T}{\partial\omega}\right|+F}\times(2.590\times10^{-3}) \qquad (8)$$

It should be noted that $\left|\dfrac{\partial T}{\partial\omega}\right|$ is not a constant but varies with the value of the control winding voltage V_c.

The variation of K and τ with V_c is easily measured. Note that application of the final value theorem to equation 6 yields

$$\omega_{ss}=\frac{AV_c}{\left|\dfrac{\partial T}{\partial\omega}\right|+F} \qquad (9)$$

and from this

$$K=\frac{A}{\left|\dfrac{\partial T}{\partial\omega}\right|+F}=\frac{\omega_{ss}}{V_c} \qquad (10)$$

Thus, the application of a number of constant magnitude control field voltages with measurement of the resulting steady-state speed ω_{ss} provides a curve of K versus V_c, as shown in Fig. 2, curve a. The dimensions of K are radians per second per volt.

In like manner, a number of step-function tests applying various values of V_c and recording the velocity-versus-time curve provide a curve of τ versus V_c, as shown in Fig. 2, curve c. To measure τ by step-function methods requires some type of tachometric device. If a tachometer is a part of the normal load assembly, there is no difficulty, but in other cases, the addition of a tachometer for measurement purposes may cause an undesired loading effect. In such cases it is more convenient to calculate the value of τ, and a method is proposed in the following.

Since A is known from stalled-torque data, and K has been determined by the steady-state speed test, the denominator of equation 7 is determined by

$$\left|\frac{\partial T}{\partial\omega}\right|+F=\frac{A}{K} \qquad (11)$$

This is also the denominator of equation 8 so that determination of the inertia J permits evaluation of τ. The results of such calculations are shown in curve b of Fig. 2. A method of measuring J, for motor alone or for motor with load, is given in the Appendix.

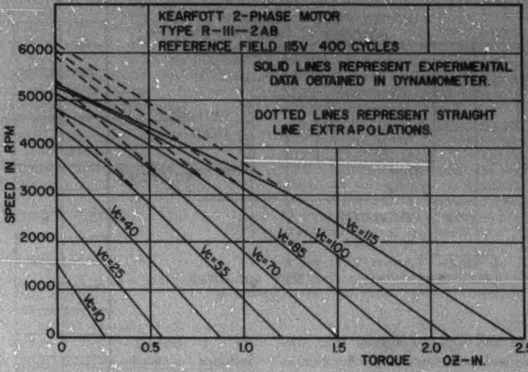


Fig. 1 (above). Speed-versus-torque curves of a 2-phase motor as determined from dynamometer tests

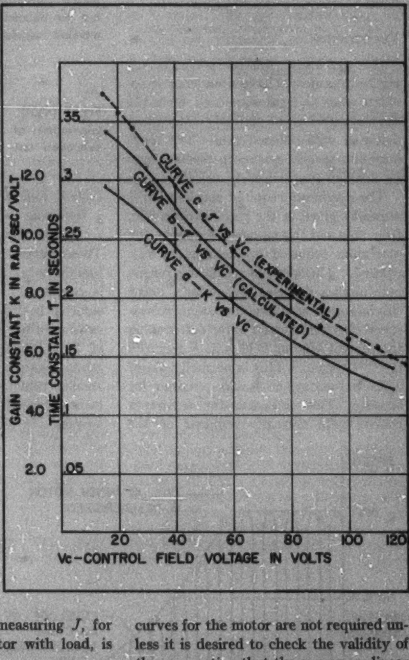


Fig. 2 (right). Curves of time constant and motor gain versus control winding voltage

Summary of Measurements Required for Curves of Fig. 2

1. Stalled torque versus V_c
2. Steady-state speed versus V_c
3. Speed-versus-time curves for a step of V_c, or measurement of J as shown in the Appendix.

Note that the torque-versus-speed curves for the motor are not required unless it is desired to check the validity of the assumption that the curves are linear below 50 per cent of rated speed. It should also be noted that J may be measured with the load geared to the motor.

Application of the Curves to the Analysis and Design of Servo Systems

In the following it is assumed that the basic specifications for the system are

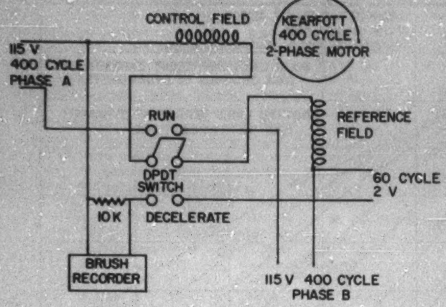


Fig. 3 (above). Switching circuit used in deceleration test for determining moment of inertia

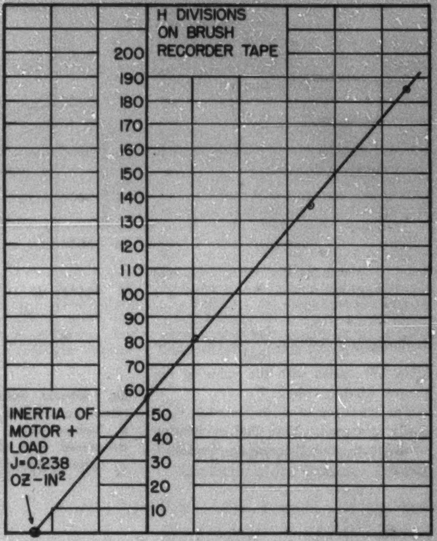


Fig. 5 (right). Extrapolation of deceleration test data

known and also the system time constants and gains caused by other components.

Verification of Stability

It is well known that servo systems using 2-phase motors have a tendency to oscillate when high gains are used, while the same system may be perfectly stable when operated with reduced gain. The tendency of a specific system to oscillate may be checked approximately as follows.

The maximum required acceleration is normally given in the specifications and, from this and the moment of inertia, the maximum required torque may be calculated. The speed at which this torque is required is also usually specified. Using this information, the motor torque-versus-speed curve is referred to for determining the control-winding field voltage required for this torque. This is normally about 1/3 rated voltage to provide a margin for slewing. This voltage is the maximum control field normally required of the system and will be called $V_{c\ \max}$.

Referring to Fig. 2, the values of K and τ corresponding to $V_{c\ \max}$ are determined. These values are inserted in the transfer-function equation and the system stability is computed. If the system is deemed sufficiently stable, it will probably operate satisfactorily. Note that for larger values of $V_{c\ \max}$, both K and τ decrease which tends toward a more stable system. For smaller values of $V_{c\ \max}$, both K and τ increase, which tends toward a less stable system. In normal operation, the error voltage approaches zero and, thus, V_c decreases from some large value toward zero. During this process, K and τ are increasing, and it is possible that a system which appears stable from the aforementioned analysis may become unstable as the system approaches correspondence. This would result in a perturbation of low amplitude.

For a system exhibiting such characteristics, the amplitude and the frequency of the perturbation may be determined approximately by applying the aforementioned method for successively smaller values of V_c until the values of K and τ will define a system which is at the stability limit. Note that the perturbation thus defined cannot increase in amplitude because any increase changes the values of K and τ in the proper direction to make the system stable.

Velocity Lag Error

If the application is such that constant velocity operation may be required, the velocity lag error will be important. The maximum anticipated velocity would be specified, and the control field voltage required to obtain this speed should be available from the data obtained in determining curve a in Fig. 2. Using this value of V_c and the curves of Fig. 2, the constants K and τ are evaluated and inserted in the transfer-function equation and, from this, the velocity lag error may

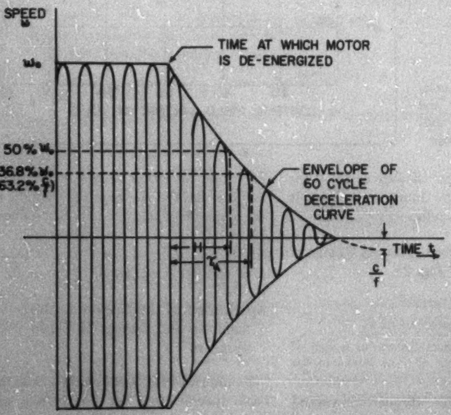


Fig. 4. Deceleration curve used in determining moment of inertia

be evaluated. Note that this computation gives the maximum velocity lag error since lower velocities require a smaller V_c which, in turn, provides a larger K, and thus the velocity lag error decreases more rapidly for reduced speeds than would be possible in the case of a linear system.

EVALUATION OF DYNAMIC ERROR

Maximum dynamic error will normally occur when maximum acceleration is required. As previously indicated, the maximum acceleration defines a value of control voltage $V_{c\,\max}$, which may be exceeded during slewing, and this value of control field voltage defines values of K and τ. These values, in turn, permit evaluation of the approximate magnitude of the dynamic error using the methods which are indicated in references 3 and 4.

Appendix. Determination of Inertia by the Deceleration Test

The object of this test is the measurement of the total inertia of the assembled motor, gear train, and load. The 2-phase motor is brought up to an initial speed ω_0 and the windings are de-energized allowing the motor to coast to a stop. If a very low voltage, about 1 or 2 volts, is impressed on the reference field, the voltage output of the control field will be proportional to the motor speed. If 60 cycles a-c are used, even for 400-cycle motors, this output can be put on a Brush recorder. A circuit for accomplishing this is illustrated in Fig. 3. It was shown by dynamometer tests that this small voltage has negligible damping effect on the motor.

The torque equation of this de-energized motor is

$$J\frac{d^2\theta_0}{dt^2}+f\frac{d\theta_0}{dt}+C=0 \tag{12}$$

where

J = total inertia of the assembled motor, gear train, and load
C = coulomb friction
f = total viscous damping of the assembled motor, gear train, and load

It should be noted that $f>F$, as the viscous damping of the motor, is included in the f term.

The instantaneous speed can be determined as

$$\omega=\left(\omega_0+\frac{C}{f}\right)\epsilon^{-\frac{f}{J}t}-\frac{C}{f} \tag{13}$$

and a plot of speed versus time is shown in Fig. 4.

The time constant of this deceleration curve is $\tau_d = J/f$. By adding a series of inertia disks J', the time constant will be changed to $\tau_d' = (J+J')/f$. These disks should be added directly on the motor shaft if possible. For this investigation, the series of inertia disks were bored to fit over the hub of the driving gear of the motor. A plot of τ_d' versus J' will be a straight line whose extrapolation to the horizontal axis will yield the desired value J.

It might be difficult to determine the time constants τ_d' of the different deceleration curves. However, if the initial speeds ω_0 of all the deceleration tests are the same, it would not be necessary to determine the time constants. The time H for the speed to decay to, say, 50 per cent of ω_0 would be proportional to τ_d' and would serve just as well. Fig. 5 is a plot of H in scale divisions of the Brush recorder tape versus J'. It will be noted that the (ω_0+C/f) ordinates of all the deceleration curves are identical.

References

1. SOME SERVO SYSTEM SINUSOIDAL STUDIES, C. F. White. *NRL Report R-3303*, Naval Research Laboratory, Washington, D. C., June 22, 1948, pp. 2–4.

2. A SPECIAL DYNAMOMETER FOR TESTING SMALL MOTORS, S. H. Van Wambeck, W. A. Stein. *Electrical Engineering*, vol. 71, June 1952, pp. 549–51.

3. SERVOMECHANISM ANALYSIS (book), G. Thaler, R. Brown. McGraw-Hill Book Company, Inc., New York, N. Y., 1953, pp. 252–54.

4. SERVOMECHANISMS AND REGULATING SYSTEM DESIGN (book), H. Chestnut, R. Mayer. John Wiley and Sons, Inc., New York, N. Y., 1951, pp. 218–20.

◆

Discussion

Kenneth A. Stone (University of Michigan, Ann Arbor, Mich.): The method described in this paper for the determination of inertia was quite interesting. Although the method is not a new one, the procedure used to determine the speed during deceleration was rather unique.

In the analysis of the servomotor transfer function, the motor damping coefficient is described as $|[(\partial T)/(\partial\omega)]|$. By limiting the discussion to the region of positive torque and speed, this does not cause any difficulty. However, inspection of a more complete set of speed-versus-torque characteristics would probably show that $(\partial T)/(\partial\omega)$ changes sign, in which case the analysis would not be valid. Certainly, operation of the motor outside of the range considered is quite possible so that the restrictions imposed by $|[(\partial T)/(\partial\omega)]|$ are not reasonable. In this respect it would be interesting to see the variation of K and τ with voltage at various negative speeds and positive speed where the curves are not assumed to be straight lines.

The large change in the time constant is of interest because one would not as intuitively expect this to occur as one might expect the change in gain. This information should be of significant value to the design engineer who works with linear analysis and the necessary assumptions.

W. A. Stein and G. J. Thaler: The authors wish to thank Mr. Stone for his comments.

It must be remembered that the paper presents a method of evaluating the transfer function of a 2-phase servomotor not coupled to a tachometer and without the necessity of obtaining dynamometer curves. In the majority of its applications, the servomotor is not required to run above 50 per cent of no-load speed. Within this prescribed limitation, the method aids the designer to bracket the system performance by established linear theory.

In dealing with the motor in the higher speed ranges or even at negative speeds, the variance of K and τ could be approximated from dynamometer curves. However, this data would have to be handled in the form of a describing function, and the servo system would have to be analyzed by nonlinear techniques.

Applications and Industry Index for 1954

1. Technical Subject Index

A-C Generators, Characteristics of Aircraft. Stratton, Matsch........165–9
A-C Generators, Transient Characteristics of Aircraft. Holloway........187–90
A-C Motors and Controls in Refineries and Pipe Lines, Application of Large. Oscarson........296–301
A-C Power System at the Fairless Works. Watkins, Derr, Fountain, Squires........343–52
A-C Systems, A Method of Calculating Current-Limiter and Fuse-Clearing Times in. Caldwell, Jensen........263–7; disc. 267
A-C Systems, Overvoltage Detection in Single and Multigenerator Aircraft. Tucker, Trbovich........420–4; disc. 424
Accurate Control of Relative Speed and Cut in a Continuous Process Line. Anger, Pettit........485–90; disc. 490
Acid Batteries, Some Discharge Characteristics of Lead. Hoxie........17–22
Acid Storage Batteries, Some Aspects of the Charge and Discharge Processes in Lead. Craig, Hamer........22–34
Additions to z-Transformation Theory for Sampled-Data Systems. Lago........403–08
Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled Aircraft Generators, Ten Years of Progress in Predicting the. Scott........455–60; disc. 460
Air-Borne Stabilized Camera Mount, Development of an. Miller, Alexander........184–6
Air Ionization as an Environment Factor. Beckett........161–5
(Air Transportation) A High-Power Servo Analyzer. Gorrill, Walley........338–43
(Air Transportation) A Method of Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems. Caldwell, Jensen........263–7; disc. 267
(Air Transportation) A Self-Excited Induction Generator with Regulated Voltage. McConnell........288–94; disc. 294
(Air Transportation) Calculations on Voltage Unbalance for 3-Phase Synchronous Systems. Wilson, Gardner........426–37; disc. 437
(Air Transportation) Carbon-Pile Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation. Richards........357–63
(Air Transportation) Correlation of Rating Data for Rotating Electric Machinery. Naumann........443–51; disc. 451
(Air Transportation) Development of a Large Ratio-Release Circuit-Breaker Mechanism. Scarcelli, Steiner........512–17; disc. 517
(Air Transportation) Environmental Testing of Small Electronic Components at High Sound Intensities. Mintz, Levine........503–09; disc. 509
(Air Transportation) Evaluation of Designs for Intermittently Heated Surfaces. Dahm, Holloway........198–205
(Air Transportation) Performance of a Constant-Speed Drive. Giloy........179–84
(Air Transportation) The Nature of Voltage Ripple on D-C Generators. Collamore........390–3
Aircraft A-C Generators, Characteristics of. Stratton, Matsch........165–9
Aircraft A-C Generators, Considerations Applicable to Automatic Paralleling of. Powell, Giloy........258–63; disc. 263
Aircraft A-C Generators, Transient Characteristics of. Holloway........187–90
Aircraft A-C Systems, Overvoltage Detection in Single and Multigenerator. Tucker, Trbovich........420–4; disc. 424
Aircraft Electric Power Systems, Economic Factors for. Bergslien, Stratton, Finison........270–9; disc. 278
Aircraft Electric System, Study of Transformerless Rectified Higher Voltage D-C. Dallas, Reising........253–7; disc. 257
Aircraft Electric Systems, A Weight Analysis of Modern. Garbarino, Hawkes, Granath........463–8; disc. 468
Aircraft Electric Systems, Sensitivity Requirements of Reactive Load Division Circuits in. Sherrard........279–86; disc. 286
Aircraft Equipment, Reliability Improvement—A Plan to Achieve and Measure Reliability of. Luckman........378–83
Aircraft Generators, Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of Blast-Cooled. Scott........455–60; disc. 460
Aircraft Generators, Thermal Developments in. Martin........493–502; disc. 502
Aircraft Inverters, High Altitude. Franklin........363–72
Aircraft Power Distribution Circuits, Experimental Determination of 400-Cycle Impedance of Wire in. Andrew........469–77; disc. 477
Aircraft Switch Testing, Comments on. Stuelpnagel, Dallas........301–05
Aircraft Tachometer Indicator—An Analysis of Design Factors Affecting Starting Performance. Akeley........398–403
Aircraft Windshields Heated by Means of Transparent Conductive Films. Ward........408–16
Altitude Aircraft Inverters, High. Franklin........363–72
Analysis of Automatic Control Systems with Nonlinear Gain Elements, Phase-Plane. Kalman........383–90
Analysis of Design Factors Affecting Starting Performance, Aircraft Tachometer Indicator—An. Akeley........398–403
Analysis of Modern Aircraft Electric Systems, A Weight. Garbarino, Hawkes, Granath........463–8; disc. 468
Analysis of Sampled-Data Control Systems, Transient. Johnson, Lindorff........147–53
Analyzer, A High-Power Servo. Gorrill, Walley........338–43
Application of Large A-C Motors and Controls in Refineries and Pipe Lines. Oscarson........296–301
Application of Short-Time Memory Devices to Compensator Design, The. Ford, Calvert........88–93; disc. 93
Automatic Control Systems with Nonlinear Gain Elements, Phase-Plane Analysis of. Kalman........383–90
Automatic Paralleling of Aircraft A-C Generators, Considerations Applicable to. Powell, Giloy........258–63; disc. 263
"Automation"—Electrolytic Tinning, Trends in. Gravenstreter, Layton........97–100; disc. 100

B

Batteries, Some Aspects of the Charge and Discharge Processes in Lead-Acid Storage. Craig, Hamer........22–34
Batteries, Some Discharge Characteristics of Lead Acid. Hoxie........17–22
Blast-Cooled Aircraft Generators, Ten Years of Progress in Predicting the Aerodynamic, Thermodynamic, and Output Characteristics of. Scott........455–60; disc. 460
Blooming Mill, Fairless Works—Electric Equipment for Slabbing Mill and. Wright, Kincaid........141–3; disc. 143
Bus Bars Under Unbalanced Load Conditions, Current Distribution in Paired-Phase. Cataldo, Shackman........438–42; disc. 442

C

Calculating Current-Limiter and Fuse-Clearing Times in A-C Systems, A Method of. Caldwell, Jensen........263–7; disc. 267
Calculations on Voltage Unbalance for 3-Phase Synchronous Systems. Wilson, Gardner........426–37; disc. 437
Camera Mount, Development of an Airborne Stabilized. Miller, Alexander........184–6
Carbon-Pile Regulator Theory: Calibration, Adjustment, and Factors Affecting Its Operation. Richards........357–63
Cell Lines, Magnetic Amplifiers in Metering Direct Current on Electrolytic. Downing........93–6; disc. 96
Characteristics of Aircraft A-C Generators. Stratton, Matsch........165–9
Characteristics of Aircraft A-C Generators, Transient. Holloway........187–90
Characteristics of Lead Acid Batteries, Some Discharge. Hoxie........17–22
Charge and Discharge Processes in Lead-Acid Storage Batteries, Some Aspects of the. Craig, Hamer........22–34
(Chemical) Some Discharge Characteristics of Lead Acid Batteries. Hoxie........17–22
Circuit-Breaker Mechanism, Development of a Large Ratio-Release. Scarcelli, Steiner........512–17; disc. 517
Circuit Design, Some Fundamentals of Equipment-Grounding. Kaufmann........227–31; disc. 231
Circuit Systems, Iron Conduit Impedance Effects in Ground. Bisson, Rochau........104–06; disc. 107
Circuits in Aircraft Electric Systems, Sensitivity Requirements of Reactive Load Division. Sherrard........279–86; disc. 286
Closed-Loop Frequency Response of Feedback Control Systems, An Extension of the Root Locus Method to Obtain. Jackson........176–9
Closed-Loop Frequency Response of Nonlinear Systems, A Generalized Method for Determining the. Prince........217–24; disc. 224
Comments on Aircraft Switch Testing. Stuelpnagel, Dallas........301–05
Commercial Buildings, Progress in Power System Engineering for. Kurt, Beeman........313–20; disc. 335
Commercial Buildings, The Relative Feasibility of 460-Volt or 208-Volt Service in. Barnett, Zimmerman, Lokay........306–13; disc. 337
(Commercial) Service Voltage Spread and Its Effect on Utilization Equipment. Barnett, Lawrence........320–5; disc. 336
Compensation of Saturating Servomechanisms, Linear. Burnett, Kendall........6–10
Compensator Design, The Application of Short-Time Memory Devices to. Ford, Calvert........88–93; disc. 93
Components at High Sound Intensities, Environmental Testing of Small Electronic. Mintz, Levine........503–09; disc. 509
Conductive Films, Aircraft Windshields Heated by Means of Transparent. Ward........408–16
Conduit Impedance Effects in Ground Circuit Systems, Iron. Bisson, Rochau........104–06; disc. 107
Considerations Applicable to Automatic Paralleling of Aircraft A-C Generators. Powell, Giloy........258–63; disc. 263
Considerations in Applying Rectifiers as a Power Supply for Hot Strip Mills. Zins, Cham........65–9; disc. 69
Constant-Speed Drive, Performance of a. Giloy........179–84
Contactor Servomechanisms, An Investigation of the Switching Criteria for Higher Order. Bogner, Kazda........118–26; disc. 126
Contactor Servomechanisms Employing Sampled Data. Chow........51–62; disc. 62
Control for Diesel-Electric Locomotives, Static Excitation. McElhenny, Smith........158–60
Control of Radio-Frequency Generators, Magnetic Amplifier. Mohr, Lee........374–8
Control of Relative Speed and Cut in a Continuous Process Line, Accurate. Anger, Pettit........485–90; disc. 490
Control Systems, Transient Analysis of Sampled-Data. Johnson, Lindorff........147–53
Control Systems with Nonlinear Gain Elements, Phase-Plane Analysis of Automatic. Kalman........383–90
Co-ordinated Fuse Protection for Low-Voltage Distribution Systems in Industrial Plants. Lebens........77–80; disc. 80
Co-ordination Considerations, Rectifier Motive Power—Inductive. King, Gordon, Hibbard........107–14; disc. 114
Correlation of Rating Data for Rotating Electric Machinery. Naumann........443–51; disc. 451
Criteria for Higher Order Contactor Servomechanisms, An Investigation of the Switching. Bogner, Kazda........118–26; disc. 126
Criteria for Servomechanism Performance, The Influence of Time Scale and Gain on. Graham, Lathrop........153–7; disc. 157
Current Distribution in Paired-Phase Bus Bars Under Unbalanced Load Conditions. Cataldo, Shackman........438–42; disc. 442
Current-Limiter and Fuse-Clearing Times in A-C Systems, A Method of Calculating. Caldwell, Jensen........263–7; disc. 267
Cut in a Continuous Process Line, Accurate Control of Relative Speed and. Anger, Pettit........485–90; disc. 490